PPP May not Hold Afterall: A Further Investigation

Serena Ng * Pierre Perron †

May 1999

Abstract

In a recent paper, Engel (1999b) presents monte-carlo evidence to suggest that unit root tests cannot detect a non-stationary component in the real exchange rate even when this component accounts for almost half of its long-horizon forecast error variance. This hidden non-stationary component led Engel to conclude that long run PPP might not hold afterall. In this note, we first highlight the extreme properties of the simulated data being considered, but concur that even when the data have less extreme properties, unit root tests will over-reject. However, the size problem can be alleviated with suitable construction of the tests. We discuss in layman's terms what steps a practitioner can take to minimize Type I error in cases when the non-stationary component is hard to detect. We also show that the contribution of the non-stationary component to long horizon forecast errors is substantially smaller than Engel reported. The key difference is that our estimates are based on forecast errors of the real exchange rate directly, rather than the forecast errors of the two components underlying it. Real exchange rate data for 19 countries are examined and estimates are obtained for the duration of the real exchange rate shocks.

*Department of Economics, Boston College. Email: serena.ng@bc.edu

†Department of Economics, Boston University and C.R.D.E., Université de Montréal. Email: perron@bu.edu

Correspondence: Serena Ng, Department of Economics, Boston College, Chestnut Hill, MA 02467.

1 Motivation

The law of one price (PPP) posits that the price level of traded goods converted to a common currency should be equal as a result of arbitrage. It would then seem natural to conjecture that national price levels converted to a common currency– the real exchange rate- should also tend towards parity. Let $p = (1-\psi)p^T + \psi p^N$ be the logarithm of the national price level of a home country. It is a geometric weighted average of the log price of traded and non-traded goods, p^T and p^N, where ψ is the share of non-traded goods. If we denote variables for the foreign country with an asterik (*), and let s_t be the log of the nominal exchange rate, then the log real exchange rate is

$$\begin{aligned} q_t &= s_t + p_t^* - p_t \\ &= x_t + y_t, \qquad (1) \\ where \qquad & x_t = s_t + p_t^{T*} - p_t^T, \\ & y_t = \psi^*[p_t^{N*} - p_t^{T*}] - \psi[p_t^N - p_t^T]. \end{aligned}$$

The real exchange rate thus has two components: a traded-goods component x_t, and a component y_t which captures the bilateral difference between the relative price of traded to non-traded goods. The real exchange rate is stationary if x_t and y_t are both stationary, or x_t and y_t are both non-stationary but that the two series have common stochastic trends. Given our prior that PPP should hold for traded goods, at least in the long run, non-stationarity of x_t is difficult to admit. Stationarity of the real exchange rate would then seem to rest on stationarity of y_t. Is this necessary? And does it matter if y_t is non-stationary?

Real exchange rates are generally found to be highly persistent. There is little dispute about this. More difficult to ascertain is whether this persistence is strong enough to be deemed non-stationary. Establishing this as a fact turns out to be a non-trivial task. As surveyed in Froot and Rogoff (1995), while early research suggests the presence of a unit root, the recent evidence supports stationarity. Many put the blame for this ambiguity on the low power of unit root tests and the lack of data with a long enough span. The views expressed in Froot and Rogoff (1995) are representative of the latest thinking. The size of unit root tests was rarely brought up as an issue.

In a recent paper, Engel (1999b) argued that root tests would fail to identify a non-stationary component in the real exchange rate even if there was one. Engel (1999b) Engel used a macroeconomic model with eight parameters to calibrate the U.S.-U.K. real exchange rate. The data support x_t as a stationary AR(1) process but suggest y_t has a unit root, implying that $q_t = x_t + y_t$ should have a unit root. But in monte carlo experiments in which q_t was tested for a unit root the way an “average person”would, Engel found that unit root tests would overwhelmingly reject the null hypothesis even though q_t has a unit root component by construction, and that the long horizon forecast error variance of y_t accounts for almost half the combined forecast error variance of x_t and y_t. Rejection rates were close to 100 percent when asymptotic critical values at the 5% level were used. So why do unit root tests reject non-stationarity?

Apart from clarifying the results of Engel, the goal of this article is to help understand why testing for a unit root in some data is so difficult and to stress that in spite of this difficulty, the ”average person” has not been doing the best that he can, both in terms of minimizing Type I error (rejecting the null hypothesis of non-stationarity) and maximizing the power of the tests. We analyze data on real exchange rates and present evidence for the life of shocks to real exchange rates using estimates of autoregressive parameters that have better properties than conventional least squares estimation.

2 The Negative Moving Average Component and the Real Exchange Rate

Consider a series $\widetilde{z}_t = z_t - \mu_t, t = 1, \ldots T$ where μ_t is a deterministic trend function which we assume is known for the moment. Suppose the (demeaned or detrended) series $\widetilde{z}_t$ is generated by an unobserved component model:

$$\begin{aligned} \widetilde{z}_t &= \tau_t + \eta_t, \\ \tau_t &= \alpha \tau_{t-1} + v_t, \\ \text{Then} \qquad \Delta \widetilde{z}_t &= (\alpha - 1)\widetilde{z}_{t-1} + e_t, \\ & e_t = u_t + \theta u_{t-1}, \end{aligned} \tag{2}$$

with θ satisfying

$$\frac{\theta}{1+\theta^2} = \frac{-\alpha\sigma_\eta^2}{\sigma_v^2 + (1+\alpha^2)\sigma_\eta^2}.$$

Without loss of generality, assume that η_t and v_t are i.i.d. and mutually uncorrelated. Suppose $\alpha = 1$ and σ_v^2 (the innovation variance to τ_t) is infinite. Then $\theta = 0$ and $\widetilde{z}_t$ is completely dominated by the random walk component. At the other extreme when $\sigma_v^2 = 0$, then $\theta = -1$ and $\widetilde{z}_t = u_t$ is i.i.d. in view of the common factor between the moving average and the autoregressive polynomial (the unit root). In between the two extremes, $\widetilde{z}_t$ is fundamentally non-stationary but also has a tendency to revert to mean. This force for mean reversion is larger the closer is θ to -1. It is this tension between non-stationarity and mean reversion that poses problems for unit root tests. The size problem arises because $\widetilde{z}_t$ behaves like a stationary process and unit root tests are fooled.

Nabeya and Perron (1994) referred to these as nearly integrated nearly white noise processes. Cases when the innovation variances are of comparable magnitudes but α is near but not exactly unity are also a problem for unit root tests. But the problem there is low power and the issue should be kept distinct from the size problem arising from a near common factor in the moving-average and the autoregressive polynomial that is being discussed here.

The size problem in testing for a unit root when there is a large negative moving average component was documented in Phillips and Perron (1988) and highlighted by Schwert (1989), among many others. To see the nature of the problem, rewrite (2) in the form of a $k^{th}-$ order augmented autoregression in $\Delta\tilde{z}_t$:

$$\begin{aligned}
\Delta\tilde{z}_t &= (\alpha-1)\sum_{i=0}^{k}(-\theta)^i\tilde{z}_{t-i-1} - \sum_{i=1}^{k}(-\theta)^i\Delta\tilde{z}_{t-i} + +e_t - (-\theta)^{k+1}e_{t-k-1}. \\
&= \beta_0\tilde{z}_{t-1} + \sum_{i=1}^{k}\beta_i\Delta\tilde{z}_{t-i} + e_{tk}, \qquad (3) \\
e_{tk} &= e_t - (\theta)^{k+1}e_{t-k-1} - \sum_{i=k+1}^{\infty}(-\theta)^i\Delta\tilde{z}_{t-i} + (\alpha-1)\sum_{i=1}^{k}(-\theta)\tilde{z}_{t-i-1},
\end{aligned}$$

where $\beta_0 = (\alpha-1)$, $\beta_i = -(-\theta)^i$. Notice that the truncation lag k plays a crucial role in the dynamic properties of e_{tk}. When θ is large and negative, lags of $\Delta\tilde{z}_t$ will have non-negligible weights in e_{tk} at large k even when $\alpha = 1$. If $\theta = -.8$, for example, we need $k > 20$ for $(-\theta)^k$ to be less than .01. Because $\Delta\tilde{z}_t$ is serially correlated, e_{tk} can be strongly serially correlated if k is small and $\theta < 0$. The severity of this problem is specific to negative values of θ because when θ is positive, $(-\theta)^i$ alternates in sign and successive lags of $\Delta\tilde{z}_{t-i}$ offset each other.

The size problem in (perhaps all) unit root tests when θ is negative can be traced to the fact that β_0 cannot be precisely estimated from (3). Nabeya and Perron (1994) and Perron (1996) analyzed the problem for the case with $k = 0$. The more general case which allows k to increase with the sample size was analyzed in Ng and Perron (1995, 1997, 1998) and Perron and Ng (1996, 1998). In those cases, $\beta_0 + 1 \equiv \alpha$ is the sum of the coefficients of an AR(k+1) model in the levels of $\tilde{z}_t$; it is this sum that is not precisely estimated.

Let us return to the real exchange rate problem. Suppose

$$y_t = y_{t-1} + w_t,$$
$$x_t = \phi x_{t-1} + m_t.$$

where w_t and m_t are i.i.d. with variance σ_w^2 and σ_m^2 respectively, and covariance $\sigma_{wm}^2 \neq 0$. Engel (1999b) offered a three equation model for exchange rate determination (reproduced in the

Appendix) for which x_t and y_t have the above time series properties. Using quarterly U.S./U.K. data over the sample 1970-1995, Engel estimated the parameters and used them to simulate 400 data points to mimic a 100 year sample. He constructed a battery of unit root tests. When estimation of an autoregression such as (3) was required, k was set to a maximum (hereafter denoted $kmax$) of 12 and a χ^2 test was then used to test for the significance of the last lag.[1] Engel showed huge size distortions in unit root tests for both the baseline parameters and for small perturbations around them. The MZ_α test developed in Perron and Ng (1996) to be more robust to size distortions when θ is negative[2] did not work as it should, and tests for the null hypothesis of stationarity did not seem immuned to the size problem. Engel also evaluated

$$R_0(h) = \frac{mse(y_{T+h} - y_{T+h|T})}{mse(x_{T+h} - x_{T+h|T}) + mse(y_{T+h} - y_{T+h|T})}, \tag{4}$$

where $mse(\cdot)$ is the mean-squared forecast error function and h is the forecast horizon.[3] For the stochastic processes assumed for x_t and y_t,

$$R_0(h) = \frac{h \cdot \sigma_w^2}{h \cdot \sigma_w^2 + \frac{(1-\phi^{2h}}{(1-\rho^2)}\sigma_m^2 + 2\frac{(1-\rho^h)}{(1-\rho^2)}\sigma_{wm}^2}.$$

Engel interprets $R_0(h)$ as the importance of the random walk component at horizon h. With $h = 400$ and assuming σ_{wm}^2 is negligible, this ratio was reported to be around .4 for the base case. Thus, variations in the non-stationary component seems important at long horizons. In spite of this, unit root tests reject non-stationarity and tests for stationarity cannot reject that null hypothesis. It appears that unit root tests have indeed missed a non-negligible permanent component badly.

Recall from the unobserved components model that the problem of a small innovation variance in the random walk component maps into a large negative moving-average component in the observed series.[4] If y_t is a random walk and x_t is a stationary AR(1), then for $q_t = x_t + y_t$,

$$\begin{aligned} \Delta q_t &= \phi \Delta q_{t-1} + e_t \\ e_t &= u_t + \theta u_{t-1}, \end{aligned} \tag{5}$$

Equation (5) is a special case of (2) with $\phi = 0$, and the key parameter is once again θ. It is related to the parameters of the processes for x_t and y_t as follows:

$$\frac{\theta}{1+\theta^2} = \frac{-\phi\sigma_w^2 - \sigma_m^2 - (1+\phi)\sigma_{wm}^2}{(1+\phi^2)\sigma_w^2 + 2\sigma_m^2 + 2(1+\phi)\sigma_{wm}^2} \tag{6}$$

[1]This is a small variation to the t test considered in Ng and Perron (1995).

[2]Engel referred to this as the PN test.

[3]Under optimal prediction, this is simply the forecast error variance and these terminologies will be used interchangeably.

[4]Ng and Perron (1997) used a similar framework to analyze the inflation series.

For the base case, $\sigma_w^2 = .328 \times 10^{-4}$ and $\sigma_m^2 = .2667 \times 10^{-2}$ with σ_{wm}^2 very small. Since σ_w^2 is 100 times smaller than σ_m^2, θ should be large and negative. For the base case, Engel calculated that $\theta = -.8$.

There are two issues involved. First, does Δq_t behave like this calibrated ARMA process? This is an empirical matter which we will take up in a later section of this article. Second, suppose Δq_t indeed has this negative moving average component and Engel's suspicion of size problems is well motivated. Then the issue is why are the size distortions so extreme, and what can be done when we encounter such a type of data? We take up this latter issue in the rest of this section.

To understand what is happening, we calculate ϕ and θ for different parameterizations of the exchange rate model. Six configurations of parameters along with the implied values of ϕ and θ are given in the Appendix.[5] Cases 1 to 3 are taken from Engel. All three have small values of σ_w^2/σ_m^2. Hence, in each case, θ is very close to the unit circle. Case 1, which is Engel's base case, has $\theta = -.9911$, (rather than -.8 which Engel reported). Cases 2 and 3 have $\theta = -.9995$ and -.9827 respectively. Table 1 reports simulations based on 2500 replications of the model in Appendix A with $T = 400$. The results by and large confirm Engel's finding. Although in no case did an exact common factor occur, the problem of parameter redundancy is severe in all three cases because θ is almost on the boundary. Since $\Delta q_t = \phi \Delta q_{t-1} + e_t$ is well approximated by $q_t = \phi q_{t-1} + u_t$, in a regression of q_t on q_{t-1}, the least squares estimator tends to identify $\phi < 1$. It is thus not surprising that MZ_α and DF reject the null hypothesis of a unit root in these three cases.

Cases 4 through 6 are our own parameterizations. Case 4 involves a close common factor between ϕ and θ, but the unit root in the data is left intact. All tests have no problem detecting non-stationarity and the size of the tests are close to the nominal size of 5%. Cases 5 and 6 have θ much further away from the unit circle, and for which we would expect unit root tests not to reject the null hypothesis. But size distortions remain noticeable. Thus, even though Engel has presented size distortions for cases of parameter redundancy that are indeed extreme, the size problem with the tests used in Table 1 is genuine.

Table 1 also confirms Engel's result that, for the base case, y_t accounts for over 40% of the forecast error variance according to $R_0(h)$. But if q_t really behaves like a stationary AR(1) process with parameter $\phi < 1$, the puzzle is perhaps not so much that unit root tests over-reject, but why the random walk component explains so much of the long-horizon forecast error variance? To see this, notice that even though $\sigma_q^2 = \sigma_x^2 + \sigma_y^2 + 2\sigma_{xy}^2$ over a sample of size T, this identity does not hold for out of sample error variances. This is because the minimum mean squared forecast

[5] Case 1 is the base case of Engel. Other configurations in his Table 3 give very similar parameter values and therefore have similar size properties.

for q_t is given by an ARIMA(1,1,1) model, and not the forecast of an AR(1) process combined with a random walk. This is important because neither x_t nor y_t has the negative moving average component that is inherent in q_t. Since we are interested in the importance of y_t in forecasting q_t, it is more appropriate to consider:

$$R_1(h) = \frac{mse(y_{T+h} - y_{T+h|T})}{mse(q_{T+h} - q_{T+h|T})}. \tag{7}$$

Since

$$(1 - \phi L)\Delta q_t = (1 - \phi L)w_t + (1 - L)m_t,$$

by direct calculations,

$$mse(\Delta q_{T+h} - \Delta q_{T+h|T}) = \sigma_w^2 + \sigma_m^2 \left[1 + \frac{(1-\phi)^2(1-\phi^{2(h-1)})}{1-\phi^2}\right].$$

Hence $mse(q_{T+h} - q_{T+h|T}) = \sum_{i=1}^{h} mse(\Delta q_{T+i} - \Delta q_{T+i|T})$. It follows that

$$R_1(h) = \frac{h\sigma_w^2}{h\sigma_w^2 + h\sigma_m^2\left[1 + \frac{(1-\phi)^2}{1-\phi^2}\right] - \sigma_m^2 \frac{(1-\phi)^2(1-\phi^{2h})}{(1-\phi^2)^2}}$$

A comparison of the denominator of R_0 and R_1 makes it clear that the h-step ahead forecast error variance of q_t can be very different from the sum of the forecast error variance of x_t and and y_t. In particular, R_0 understates the mean-squared error of q_t and therefore overstates the importance of the permanent component. As we see from Table 1, R_1 is small at large h if σ_w^2/σ_m^2 is small. At h=400, R_1 is quite close to zero for the base case. The puzzle that unit root tests reject the null hypothesis even when a non-stationary component explains an important fraction of long run forecast error variance is specific to the use of R_0 in the calculations.

2.1 A Digression on Measuring Persistence

The decomposition of long-horizon forecast errors is one way to assess the importance of the permanent component. But as seen above, calculations of R_1 necessitates knowledge of the moving-average component. But given the tradition in the literature to favor autoregressive rather than mixed ARMA models, how are the alternative ways of measuring persistence? Suppose a $k+1^{th}$ order autoregression is estimated, and let $\widehat{\alpha} = \sum_{i=1}^{k+1} \widehat{\alpha}_i$, be the sum of the estimated autoregressive coefficients. Consider two measures of persistence, both aim to capture the time required for a fraction, say τ, of the effect to a unit shock is complete. Define

$$\begin{aligned} J_0 &= sup_j\ |\partial z_{t+j}/\partial u_t| \leq 1 - \tau, \\ J_1 &= log(1-\tau)/log(\widehat{\alpha}). \end{aligned}$$

When $\tau = .5$, J_0 is the period beyond which the (absolute) response to a unit shock in u_t no longer exceed .5. On the other hand, J_1 is the half life of a shock as implied by estimated sum of the autoregressive coefficients. The difference between the two is that J_0 is based on the moving-average representation of the estimated model and hence depends on all $k+1$ parameters in its autoregressive representation. In contrast, J_1 depends only on the estimated sum of the autoregressive parameters.

Our conjecture is that for data with a negative moving average component, J_1 will overstate the effects of the permanent component over short horizons. To see why this is the case, consider an ARMA(1,1) process with parameters α and θ for $\tilde{z}_t$ with $\alpha+\theta \neq 0$. From the infinite moving-average representation, we see that the response of z_{t+j} to a unit shock in period t is:

$$\partial \tilde{z}_{t+j} / \partial u_t = \alpha^{j-1}(\alpha + \theta). \tag{8}$$

While α controls the slope of the impulse response function, θ helps pin down the amplitude. When $\alpha = 1$, both statistics evaluated at τ close to one will be large, since α^i does not vanish in the case of J_0, and $log(1)$=0 in the case of J_1. Thus, both statistics will have no problem revealing that the complete adjustment to a shock will be infinitely long. But what if one's interest is in smaller values of τ (such as .5), which are associated with horizons relevant for policy analysis? Since the dynamic effects of a shock at any lag is $\alpha^{j-1}(\alpha + \theta)$, it will be much smaller than α^{j-1} when θ is negative. Thus, J_1 will overstate the duration of adjustment. The J_0 statistic utilizes the autoregressive estimates at all lags and thus provides a more reliable measure of the speed at which the effect of a shock dissipates.

An alternative to J_0 is the autocorrelation function, say, $\Gamma(j)$. Indeed, both the estimated sum of the autoregressive coefficients and $\Gamma(j)$ are widely used as measures of persistence in macroeconomic analysis.[6] The potential problem with $\Gamma(j)$ is that it is typically evaluated at only small values of j. But for processes that are both persistent and have a tendency for mean reversion, the j that is consistent with a large τ could be very large. In this regard, $\Gamma(j)$ could understate the importance of the permanent component. The J_0 statistic has the advantage that we do not need an a priori choice on j; it is endogenously determined once we pick the cut-off point, τ.

3 Testing for a Unit Root Once Again

Some, including ourselves, have argued[7] that there is always a non-stationary representation for a time series that is arbitrarily close to a stationary representation. Because of this potential for

[6] For example, Mcgrattan, Chari and Kehoe (1998) used the degree of serial correlation in the real exchange rate to judge whether a sticky price model can replicate the observed persistence in the real exchange rate. Bergin and Feenstra (1999) used the first and fourth order autocorrelation coefficients to assess the degree of stickiness in the real exchange rate.

[7] See Perron and Ng (1996), Cochrane (1991), Faust (1996), Blough (1992).

observational equivalence, any test that has high power rejecting the null hypothesis of a unit root when the signal of the non-stationary component is strong must also have a large size when this signal is week. Consider once again the mapping from the relative size of the innovation variance to θ. The near-observational equivalence problem can now be stated as follows:- when using unit root tests with asymptotic critical values, there will exist values of θ in the range $(-1, x)$ for some $-1 < x < 0$, say, such that liberal size distortions will surface. The value of x will depend on the sample size and the test used, but it will always approach -1 as the sample size increases. That is to say, the range over which size distortions occur will diminish.

Engel used an example to show that the problem of observational equivalence could indeed happen in real data. The tests he considered really ought to have supported non-stationarity, though the defense that $\theta = -.99$ is not an interesting case could perhaps be invoked. The more serious problem is that for sample sizes commonly encountered, the value of x where size distortions start to appear is not -.99, but much further away from -1. Depending on the test, x could be anywhere from -.4 to -.8 for $T = 100$. This is worrisome because there will exist empirically important time series which are genuinely non-stationary, and would yet be classified as stationary. Cases 3, 5 and 6 documented earlier are representative of the problem. If R_1 is .3 at $h = 400$ as in case 5, we would indeed want unit root tests not to reject stationarity. Table 1 shows that the rejection rate is too high. Most troublesome are cases 5 and 6.[8] Even with a sample size of 400, the statistics can reject with 80% probability instead of 5%.

How prevalent are such time series? In our experience and as we will see in the next section, variables such as inflation tend to have this property, and we are in the process of a more complete documentation of such data. While a formal test of parameter redundancy is difficult because the maximum likelihood estimates of the autoregressive and moving average parameters are not precise when there is a near common factor,[9] the symptoms are there for us to detect. From our previous work, the kernel estimate of the spectral density at frequency zero based upon $\widehat{e}_{t0}$ (i.e. the least squares residuals) should be very different from a particular autoregressive spectral estimate of e_{t0} (which does not depend on $\widehat{e}_{t0}$).[10] There should also be sharp differences between the Phillips-Perron Z tests and MZ tests even though the two differ only by a term that should vanish at rate T. The premise of our latest work is precisely to exploit such information to robustify the size of DF and the class of MZ tests. This is achieved by parameterizing the model and/or finding

[8]Case 4 may seem somewhat surprising because if the dominant root is unity, then one might expect R_1 to be larger. Furthermore investigation reveals that we need σ_w^2 to be at least 50 times larger σ_m^2 for the forecast error variance to be overwhelmed by the random walk component. Therefore, a weak force for mean reversion is still at play.

[9]See Clark (1988).

[10]These issues are discussed in Ng and Perron (1997).

estimators such that the sum of the autoregressive coefficients and the nuisance parameters can be estimated as precisely as possible. We now provide a non-technical summary of this work. All statistics considered are defined in the Appendix.[11]

To begin, recall that in the above discussion $\tilde{z}_t$ is the demeaned series. That is, $\tilde{z}_t = z_t - \mu_t$, where μ_t is a vector of deterministic components. For persistent data, least squares detrending is inefficient. Elliott, Rothenberg and Stock (1996) showed that using GLS detrended data to construct the DF statistic can yield substantial power gains. Ng and Perron (1998) showed that these power gains extend to the Z and MZ tests. As a first step, therefore, one should first quasi-transform the data at $\bar{\alpha} = 1 + \bar{c}/T$, where $\bar{c}$=-7.0 in the constant case and -13.5 in the linear trend case. Then use GLS to obtain estimates of the coefficients on the deterministic components. The discussion to follow assumes that all regressions are based on GLS detrended data and the object of interest is whether there is a unit root in the detrended data $\tilde{z}_t$.

The DF-GLS is the $t-$ statistic on β_0 in the k^{th} order augmented autoregression (3).

- The Problem: $\widehat{\beta}_0$ is biased if k is small because e_{tk} is serially correlated.
- The Fix: Select a large k when necessary.
- Implementation: Use the MAIC to select k in the augmented autoregression (3), where

$$\begin{aligned} MAIC &= Argmin_{k=0,...kmax} \ln(\widehat{\sigma}_k^2) + \frac{2(\widehat{\tau}_T(k) + k)}{T}, \\ \text{with } \widehat{\tau}_T(k) &= (\widehat{\sigma}_k^2)^{-1} \widehat{\beta}_0^2 \sum_{t=1}^{T} \widetilde{y}_{t-1}^2, \end{aligned}$$

where $\widehat{\sigma}_k^2 = T^{-1} \sum_{t=1}^{T} \widehat{e}_{tk}^2$.

The key to the fix is the selection of k. The MAIC is motivated by the observation that the bias in $\widehat{\beta}_0$ decreases non-linearly as k increases. Model selection rules such as the AIC and BIC do not take this non-linearity into account; they under-penalize models with a small k and select autoregressive approximations that are too parsimonious for models with negative θ. The MAIC explicitly accounts for the strong dependence of the bias in $\widehat{\beta}_0$ on k via the term $\tau_T(k)$. The MAIC reduces to the standard AIC when this dependence is absent (such as ARMA noise functions with autoregressive and moving average roots far from the unit circle).

[11] This has been the basis of work in Ng and Perron (1995, 1997, 1998), and Perron and Ng (1995, 1996).

The Z-GLS requires *(a),* a least squares estimate of α from the regression $\widetilde{z}_t = \alpha \widetilde{z}_{t-1} + e_{t0}$, and *(b),* the short and long run variance (the non-normalized spectral density at frequency zero) of e_{t0}.

- The Problem: *(a),* $\widehat{\alpha}$ is severely biased because e_{t0} is strongly serially correlated, and because of this, *(b),* the estimated residuals $\widehat{e}_{t0}$ provide poor estimates of both the short-run and the long-run variance.
- The Fix: Remove any dependence of Z-GLS on $\widehat{\alpha}$.
- Implementation: *(a),* Apply a correction factor to Z-GLS to give MZ-GLS; *(b),* Estimate the short run variance under the null hypothesis and the long run variance using an autoregressive spectral density estimator based upon (3); *(c),* when estimating the autoregressive spectral density, use the MAIC to select k.

The modification factor removes the direct dependence of Z-GLS on $\widehat{\alpha}$. With Z_α^{GLS}, for example, this factor is simply $(T/2)(\widehat{\alpha}-1)^2$. Note that $MZ_\alpha^{GLS} = Z_\alpha^{GLS} + \frac{T}{2}(\widehat{\alpha}-1)^2$ can be rewritten as $(T^{-1}\widetilde{z}_T^2 - s^2)(2T^{-2}\sum_{t=1}^T \widetilde{z}_{t-1}^2)^{-1}$. The removal of the dependence on $\widehat{\alpha}$ is complete since s^2 also does not depend on $\widehat{\alpha}$ if the spectral density is estimated using (3). To obtain a precise estimate of s^2, select k using the MAIC for the same reason given for the DF-GLS.

The only remaining issue is that, as in the AIC and the BIC, we need to specify a *kmax* for the MAIC. Our theoretical results only provide guidance about the rate of increase of k relative to the sample size and does not pin down a *kmax* for empirical work. Our recommendation is to use a *kmax* that varies with the sample size, such as $kmax = 12(T/100)^{1/4}$. But because the k that is required to make e_{tk} approximately serially uncorrelated depends on the data generating process, there could be cases when such a *kmax* might not be large enough. That is, *kmax* could bind. In that case, reset *kmax* to some larger number and reoptimize the MAIC function. In our experience with simulations and empirical applications, setting this *kmax* will yield substantial size improvements for θ no smaller than -.8. We have encountered few economic time series with θ even closer to the boundary. Increasing *kmax* further yields small improvements in size, but risks power loss.

Use of the old and new tests can make an important difference both in terms of size and power, especially in the case of a linear time trend. For example, when $\theta = -.8$ the DF with k selected by the BIC has a rejection rate of .5 instead of .05; with Z_α, the size is close to 100%. Using the MAIC to select k, the rejection rates for DF-GLS and MZ_α^{GLS} are .149 and .084 respectively. These are huge reductions in Type I error. As for power, the gains are also non-trivial. For an iid errors, the rejection can be increased from 30% to close to 50% even for sample size reasonably small at the alternative $\bar{\alpha} = 1 + \bar{c}/T$, where $\bar{c} = -7$ for $p = 0$ and -13.5 for $p = 1$. The contrasts are somewhat less dramatic for the constant only case but remain non-trivial.

Extensive simulations for MZ-GLS and DF-GLS have been reported elsewhere for pure AR and pure MA noise functions.[12] In Table 2, we use the real exchange rate example to highlight two points. First, even with with $\theta = -.99$ as in the base case, the probability of a Type I error with the new tests is much smaller. Second, when θ is further away from the boundary (such as case 5), the size is approximately correct.

It should be emphasized that proper implementation of the new tests is extremely important. As seen from Table 2, use of the MAIC without modifying Z_α to MZ_α^{GLS} will do little to reduce size distortions. Table 3 shows that, for the same $kmax$, the BIC always have larger size distortions; increasing $kmax$ does little to help. In the simulations, it selects $k = 1$ on average. Under the MAIC, k is, on average, 12 when $kmax = 20$. Thus, use of the modified statistics without appropriately selecting k will also be ineffective. Letting k be the default used in software packages is also undesirable.

How about the t test recommended in Ng and Perron (1995)? Let us consider case 6. From Table 1, we see that the size is still over .8. The size problem here arises not because of the t test per se, but with $kmax$. While $kmax = 12$ might be large enough for most data, it is not when there is a negative moving average component as large as -.97. Simulations show that size distortions will fall if we increase $kmax$ for the t-test. But even if we had done so, the DF and MZ in Table 1 will remain inferior to the DF-GLS and MZ-GLS of Table 2. This is because the statistics in Table 1 are not based on GLS detrended data.[13] Table 4 reports the average $\widehat{\alpha} = 1 + \widehat{\beta}_0$ for different values of k. First note that when there is no common factor (such as cases 4 and 5), adding lags neither harm nor hurt the estimate. But in some cases, $\widehat{\beta}_0$ can be made substantially more precise if we increase k (compared to Table 1). Thus, it is GLS detrending along with the selection of k that improve the precision of the estimates of sum of the autoregressive coefficients. Of the two tests, the MZ-GLS tests hold a size advantage while the DF-GLS has better power, especially for sample sizes less than 150.

As an example, consider log(GDP), inflation in GDP, and in the consumption deflator over 1962q1-1998q4 ($T = 160$). The data are taken from FRED[14]. Both the DFGLS and MZ-GLS use GLS detrending and the MAIC with kmax set to 14. We also report DF and Z_α, both based on least squares detrending and k selected using the BIC. This is perhaps the most commonly used

[12] There are two differences between those simulations and the current setting. First, in those simulations θ was invariant to the variance of e_t, but here θ here varies with σ_e^2. Second, the current setting amounts to an ARMA noise function in our previous work. But neither can account for the size discrepancies between the reasonably accurate size we found and the huge size distortions reported in Engel.

[13] The weakness of the t test is that it tends to overparameterize and induces power loss in cases when large lags are not necessary.

[14] The web site is http://www.stls.frb.org/fred.

method in practice, except that $kmax$ is higher than most practitioners use.

We first estimate an ARMA(1,1) model to obtain a rough idea of the size of the moving average component. For log GDP, $\widehat{\theta}$ is positive suggesting size distortions is not an issue for all tests. Indeed, no test rejects a unit root in GDP around a linear deterministic trend, and MAIC and BIC chooses k of 2 and 1 respectively. Consider the two inflation series. Both estimates of θ are negative, and even though the point estimates are far from -1, problems with inference already surface. The Z_α rejects a unit root in both cases, while the DFGLS and MZ-GLS do not. The DF is known to be somewhat more robust than Z_α when θ is negative, but it too rejects a unit root at the 5% level for one series and 10% level for the other. For both inflation series, the BIC selects a lag length of 0. The MAIC selects 2 and 9 respectively. This shows, first, that the BIC will not pick a large k even when $kmax$ is high[15], and second that the MAIC does not necessarily pick the largest k possible. A comparison of $\widehat{\alpha}^{GLS}$ and $\widehat{\alpha}^{OLS}$ shows that in general, the OLS/BIC combo yield lower estimates of α. Also of interest is a comparison of J_0 and J_1. Evaluating them at $\tau = 1/2$, we see that the half-life of a shock to GDP is around 35 years, with little difference between J_0 and J_1. However, when there is a moving average component, J_1 indicates a much longer half life than J_0 as we conjectured earlier. On the other hand, the largest autoregressive root is appropriate for evaluating τ closer to 1. Hence when $\tau = .8$, both J_0 and J_1 suggest that it will take about 50 quarters for 80% of the effect of the shock to dissipate. The results also reinforce the finding that all series have a unit root.

4 Back to the Real Exchange Rate

One issue that arises frequently in the analysis of the exchange rate (real or nominal) is whether combining the low volatility data before the Bretton Woods agreement with data which are more volatile after the agreement will affect the size of unit root tests.[16] This issue of a break in variance was studied by Hamori and Tokihisa (1997) for the normalized least squares estimator. The authors find that the break fraction and the relative variance in the two regimes will enter the limiting distribution of the test statistic, and in monte carlo experiments, combining the data of the two regimes will lead to over-rejections of the unit root tests. Although no formal analysis is available for other test statistics, there is little doubt that the qualitative conclusion will generalize. The focus of most exchange rate analysis on one regime is justified.

Table 5 presents estimates of θ from an ARMA(1,1) model on the nominal and the real exchange rate as measured by the consumption deflator. The data are for the period 1973q1-1997q2, taken

[15]Note in passing that the t-test will, but it does not do so judiciously and hence loose power when a large k is not necessary.

[16]See, for example, the discussion in Froot and Rogoff (1995), Section 2.3.5.

from the OECD sectoral database. Results for the GDP deflator are similar and not reported. In all cases, the moving average component is estimated to be positive with t statistics larger than 1.6 in absolute value. Recall that the Engel's basic premise was that Δq_t could have a large negative moving-average component. Evidence for this is found in only two countries, Australia and Korea. The U.K. data, which was the basis of Engel's analysis, clearly did not exhibit a negative θ. Since there is hardly any evidence of a negative θ in either q_t or Δq_t, the size of unit root tests should not be an issue. Applying the new (and old) tests, we can only reject a unit root in the real exchange rate for Canada, and only marginally.

Interest in establishing whether there is a unit root in the real exchange rate arises because we want to understand how long it will take for a shock which perturbs PPP to work itself out. And rightly argued, with 25 years of data, unit root tests may indeed have low power in providing a precise I(1)/I(0) classification. But the autoregressive coefficient estimates are still informative. Consider once again J_0 along with bootstrapped standard errors.[17] With the exception of Korea, Greece, and Portugal, the half life of real exchange shocks is between nine and fifteen quarters, in line with the consensus estimate of 4.5 years from panel studies. Evaluating τ at .8 gives a clearer picture of the relative persistence across countries. Japan has the fastest speed of adjustment, with 80% completed in 15 quarters. This is followed by Canada, Ireland, Sweden, France and Italy. Adjustments in the remaining countries take over 20 quarters to complete, with Greece and Korea being the outliers.

5 Conclusion

From an econometric perspective, Engel's conclusion that long-run PPP may not hold is valid because we fail to reject a unit root in the real exchange rate. While the size issue Engel raised is a valid methodological problem, the issue is not relevant to exchange rate data we investigated. Engel's result is likely an artifact of the imprecise estimates of the parameters used to calibrate the model. We have taken the occasion to emphasize that there are steps an "average person" can take to maximize power and minimize size distortions in unit root tests. The estimation strategy also provides more precise estimates of the autoregressive parameters which can be informative about economic dynamics. Our estimates put the half life of shocks to real exchange rate between nine and fifteen quarters, though there are more variations in the time required to complete 80% of the adjustments.

There remains the question of whether we should care if a non-stationary component in q_t exists? Such an issue is of independent interest because it is relevant whenever testing a variable

[17]Because of the lack of a negative moving-average component, estimates for J_1 are similar.

which has subcomponents (such as CPI, industrial production) is at stake. The answer depends on the objective of the exercise. Take industrial production. If an economist was asked "Are all sectors stationary?", then he should document as clearly as possible unit root tests results on all sectors. But if this economist was asked "is industrial production non-stationary", then there is no value-added in knowing if there is a permanent component in, say, the output for shoelaces. One might think otherwise if it was the production of automobiles rather than shoelaces that has a permanent component. But if the variations in automobiles are important enough, it will be reflected in industrial production anyways. In the end, unit root tests on the components are neither necessary nor sufficient for establishing a unit root in the aggregate variable. On the other hand, if we were interested in the source of the unit root in the aggregate, analysis of the components will be necessary. But establishing the existence and the source are two different questions.[18] In the case of the real exchange rate, little is lost from not knowing that a permanent component in y_t exists if all we want to know is whether the real exchange rate has a unit root.

[18] Such an analysis was provided by Engel (1999a).

Appendix

Test Statistics

The Said-Dickey-Fuller DF-GLS test due to Dickey and Fuller (1979), Said and Dickey (1984) and Elliott et al. (1996) is the t statistic on $\widehat{\beta}_0$ from the augmented autoregression:

$$\Delta\widetilde{z}_t = \beta_0 \widetilde{z}_{t-1} + \sum_{j=1}^{k} \beta_j \Delta\widetilde{z}_{t-j} + e_{tk}, \tag{9}$$

where $\widetilde{z}_t$ is $z_t - \widehat{\beta}' d_t$, $\widehat{\beta}$ is the GLS estimate of the coefficients on the deterministic terms d_t.
The Phillips-Perron test is

$$Z_\alpha = T(\widehat{\alpha} - 1) - (s^2 - s_v^2)(2T^{-2} \sum_{t=1}^{T} \widetilde{z}_{t-1}^2)^{-1},$$

where $\widetilde{z}_{t-1}$ are the residuals from a regression of z_{t-1} on d_t, $t = 1 \ldots, T$, and $\widehat{\alpha}$ is the least squares estimate from the regression

$$\widetilde{\widetilde{z}}_t = \widehat{\alpha}\widetilde{z}_{t-1} + \widehat{v}_t, \tag{10}$$

where $\widetilde{\widetilde{z}}_t$ are the residuals from a regression of z_t on d_t (the deterministic terms), $t = 1, \ldots, T$, $s_v^2 = T^{-1} \sum_{t=1}^{T} \widehat{v}_t^2$ and $t_\alpha = (\widehat{\alpha} - 1)(\sum_{t=1}^{T} \widetilde{z}_{t-1}^2)^{1/2} / s_v$ is the standard t-ratio to test the null hypothesis of a unit root. The term s^2 is a consistent estimate of the spectral density at frequency zero.
The Modified Phillips-Perron test MZ_α is

$$MZ_\alpha = \frac{T^{-1}\widetilde{z}_T^2 - s^2}{2T^{-2} \sum_{t=1}^{T} \widetilde{z}_{t-1}^2} \approx Z_\alpha + \frac{T}{2}(\widehat{\alpha} - 1)^2.$$

The autoregressive estimate of the spectral density at frequency zero of v_t, is defined as:

$$s_{AR}^2 = s_{ek}^2 / (1 - \widehat{\beta}(1))^2, \tag{11}$$

where $\widehat{\beta}(1) = \sum_{i=1}^{k} \widehat{\beta}_i$, $s_{ek}^2 = T^{-1} \sum_{t=k+1}^{T} \widehat{e}_{tk}^2$, with $\widehat{\beta}_i$ and $\{\widehat{e}_{tk}\}$ obtained from the autoregression (9).

References

Bergin, P. and Feenstra, R. (1999), Pricing to Market, Staggered Contracts, and Real Exchange Rate Persistence, NBER WP 7026.

Blough, S. (1992), The Relationship between the Power and the Level of Generic Unit Root Tests in Finite Samples, *Journal of Applied Econometrics* **7**, 295–308.

Clark, P. K. (1988), Nearly Redundant Parameters and Measures of Persistence in Economic Time Series, *Journal of Economic Dynamics and Control* **12**, 447–461.

Cochrane, J. H. (1991), A Critique of the Applicaion of Unit Root Tests, *Journal of Economic Dynamics and Control* **15**, 275–284.

Dickey, D. A. and Fuller, W. A. (1979), Distribution of the Estimators for Autoregressive Time Series with a Unit Root, *Journal of the American Statistical Association* **74**, 427–431.

Elliott, G., Rothenberg, T. J. and Stock, J. H. (1996), Efficient Tests for an Autoregressive Unit Root, *Econometrica* **64**, 813–836.

Engel, C. (1999a), Accounting for U.S. Real Exchange Rate Changes, Journal of Political Economy,forthcoming.

Engel, C. (1999b), Long Run PPP May Not Hold After All.

Faust, J. (1996), Near Observational Equivalence and the Theoretical Size Problems with Unit Root Tests, *Econometric Theory* **12**, 724–731.

Froot, K. and Rogoff, K. (1995), Perspectives on PPP and Long-Run Real Exchange Rates, *in* G. Grossman and K. Rogoff (eds), *Handboook of International Economics*, North Holland.

Hamori, S. and Tokihisa, A. (1997), Testing for a Unit Root in the Presence of a Variance Shift, *Economic Letters* **57:3**, 245–253.

Mcgrattan, E., Chari, A. and Kehoe, P. (1998), Can a Sticky Price Model Generate Volatile and Persistent Real Exchange Rates, Federal Reserve of Minneapolis Staff Report 223.

Nabeya, S. and Perron, P. (1994), Local Asymptotic Distribution Related to the AR(1) Model with Dependent Errors, *Journal of Econometrics* **62**, 229–264.

Ng, S. and Perron, P. (1995), Unit Root Tests in ARMA Models with Data Dependent Methods for the Selection of the Truncation Lag, *Journal of the American Statistical Association* **90**, 268–281.

Ng, S. and Perron, P. (1997), Estimation and Inference in Nearly Unbalanced and Nearly Cointegrated Systems, *Journal of Econometrics* **79**, 54–81.

Ng, S. and Perron, P. (1998), Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power, Boston College WP 369.

Perron, P. (1996), The Adequacy of Asymptotic Approximations in the Near-Integrated Autoregressive Model with Dependent Errors, *Journal of Econometrics* **70**, 317–350.

Perron, P. and Ng, S. (1996), Useful Modifications to Unit Root Tests with Dependent Errors and their Local Asymptotic Properties, *Review of Economic Studies* **63**, 435–465.

Perron, P. and Ng, S. (1998), An Autoregressive Spectral Density Estimator at Frequency Zero for Nonstationarity Tests, *Econometric Theory* **14**, 560–603.

Phillips, P. C. B. and Perron, P. (1988), Testing for a Unit Root in Time Series Regression, *Biometrika* **75**, 335–346.

Said, S. E. and Dickey, D. A. (1984), Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order, *Biometrika* **71**, 599–607.

Schwert, G. W. (1989), Tests for Unit Roots: A Monte Carlo Investigation, *Journal of Business and Economic Statistics* **7**, 147–160.

Engel's Model and Parameterizations

$$\begin{aligned} \Delta y_t &= a u_t, \\ \Delta s_t &= -\delta(s_t + p^{*T} - p^T) = b u_t + c v_t, \\ \Delta(p_t^T - p_t^{*T}) &= \gamma(s_t + p^{*T} - p^T) + d\epsilon_t + f v_t + g u_t, \end{aligned}$$

where u_t, v_t, ϵ_t are iid N(0,1) and are mutually uncorrelated. The model then implies

$$x_t = \phi x_{t-1} - d e_t + (c - f) v_t + (b - g) u_t,$$

with $\phi = 1 - \delta - \gamma$. Thus, $\sigma_w^2 = a^2$, $\sigma_m^2 = (c - f)^2 + (b - g)^2 + d^2$, and $\sigma_{wm}^2 = a(b - g)$.

Structural Parameters and Implied Values of ϕ and θ

Case	ϕ	θ	σ_w/σ_m^2	a	d	c	δ
1	0.9196	-0.9911	0.0123	0.00573	0.01129	0.05077	0.08038
2	0.9760	-0.9995	0.0005	0.00100	0.01890	0.04150	0.02400
3	0.9524	-0.9827	0.1416	0.01800	0.01550	0.04560	0.04761
4	0.9196	-0.9259	3.7503	0.10000	0.01129	0.05077	0.08038
5	0.8500	-0.8923	0.9376	0.05000	0.01129	0.05077	0.15000
6	0.8500	-0.9713	0.0375	0.01000	0.01129	0.05077	0.15000

In the simulations $\gamma = 0$, $b = .001088$, $f = .000632$, $d = .011286$. $\phi = 1 - \gamma - \delta$ and θ is determined according to (6). Cases 1,2 and 3 are from Tables 2 and 4 of Engel (1999b).

Table 1: Statistics k selected by the t test, $kmax = 12$.

Case	ϕ	θ	MZ_α	DF	Z_α	$\widehat{\alpha}$	$R_0(400)$	$R_1(400)$
1	0.9196	-0.9911	0.9604	0.9120	0.9664	0.9188	0.4314	0.0117
2	0.9760	-0.9995	0.4436	0.3040	0.4592	0.9653	0.0092	0.0005
3	0.9524	-0.9827	0.4476	0.3452	0.4584	0.9628	0.8404	0.1215
4	0.9196	-0.9259	0.0748	0.0584	0.0784	0.9848	0.9957	0.7827
5	0.8500	-0.8923	0.1604	0.1336	0.1672	0.9763	0.9905	0.4646
6	0.8500	-0.9713	0.8400	0.7944	0.8668	0.8955	0.8063	0.0335

ϕ and θ are parameters for $\Delta q_t = \phi\Delta q_{t-1} + e_t$, $e_t = u_t + \theta u_{t-1}$ as implied by the parameters in Table 1. The rejection rates for MZ_α, DF, and Z_α based on the 5% asymptotic critical values of -14.1, -2.86, and -14.1. The tests are based on OLS demeaned data. $R_0(400)$ and $R_1(400)$ assess the importance of the permanent component in long horizon forecast errors. These are defined as 4 and 7 in the text.

Table 2: Modified Statistics with k selected by the MAIC

Case	ϕ	θ	$Z_{\alpha^{GLS}}$			MZ_α^{GLS}			DF^{GLS}		
kmax			12	20	40	12	20	40	12	20	40
1	0.9196	-0.9911	0.9036	0.8912	0.9028	0.6928	0.5560	0.3256	0.6824	0.5468	0.2884
2	0.9760	-0.9995	0.3436	0.3556	0.3628	0.4136	0.3788	0.3344	0.3964	0.3500	0.2560
3	0.9524	-0.9827	0.3380	0.3592	0.3556	0.3028	0.2744	0.2188	0.2936	0.2468	0.1744
4	0.9196	-0.9259	0.0636	0.0608	0.0568	0.0572	0.0568	0.0708	0.0568	0.0460	0.0428
5	0.8500	-0.8923	0.0964	0.1072	0.0984	0.0828	0.0764	0.0780	0.0816	0.0656	0.0464
6	0.8500	-0.9713	0.7220	0.7092	0.7236	0.5160	0.3180	0.1360	0.5128	0.3292	0.1364

The 5% asymptotic critical value for Z_α^{GLS} and MZ_α^{GLS} is -8.1, and for DF^{GLS} is -1.91.

Table 3: Modified Statistics with k selected by the BIC

Case	ϕ	θ	$Z_{\alpha^{GLS}}$			MZ_α^{GLS}			DF^{GLS}		
kmax			12	20	40	12	20	40	12	20	40
1	0.9196	-0.9911	0.9964	0.9932	0.9948	0.8876	0.8800	0.8908	0.8880	0.8828	0.8900
2	0.9760	-0.9995	0.4216	0.4264	0.4316	0.4888	0.4900	0.4904	0.5080	0.5040	0.5068
3	0.9524	-0.9827	0.4776	0.4900	0.4788	0.4068	0.4328	0.4460	0.4100	0.4380	0.4536
4	0.9196	-0.9259	0.0748	0.0636	0.0636	0.0696	0.0692	0.0656	0.0712	0.0712	0.0696
5	0.8500	-0.8923	0.1944	0.1948	0.2044	0.1524	0.1632	0.1700	0.1600	0.1668	0.1716
6	0.8500	-0.9713	0.9808	0.9796	0.9816	0.8636	0.8464	0.8512	0.8628	0.8488	0.8480

Table 4: Estimates of $\alpha \equiv 1 + \beta_0$ for different $kmax$ and Model Selection Strategies

Case	ϕ	θ	k=0	maic(12)	maic(20)	maic(40)	baic(12)	baic(20)	baic(40)
1	0.9196	-0.9911	0.9188	0.9550	0.9586	0.9648	0.9469	0.9472	0.9471
2	0.9760	-0.9995	0.9653	0.9784	0.9790	0.9804	0.9768	0.9768	0.9769
3	0.9524	-0.9827	0.9628	0.9815	0.9821	0.9841	0.9789	0.9783	0.9781
4	0.9196	-0.9259	0.9848	0.9943	0.9946	0.9949	0.9939	0.9938	0.9938
5	0.8500	-0.8923	0.9763	0.9918	0.9924	0.9930	0.9892	0.9893	0.9890
6	0.8500	-0.9713	0.8955	0.9529	0.9615	0.9706	0.9344	0.9366	0.9361

Table 5: Estimates of θ for Nominal and Real Exchange Rate: 1973:1-1997:2

	s_t	Δs_t	q_t^c	Δq_t^c
aus	0.193	-0.644	0.216	-0.745
aut	0.380	-0.790	0.354	0.593
can	0.426	-0.432	0.425	0.916
che	0.367	-0.677	0.358	0.764
deu	0.383	-0.464	0.361	0.677
dnk	0.363	-0.989	0.375	0.459
fra	0.455	-0.331	0.443	0.478
gbr	0.245	-0.406	0.231	0.324*
grc	0.192	0.975	0.182	0.248*
ire	0.310	-0.650	0.290	0.423*
ita	0.412	0.754	0.423	0.740
jpn	0.399	-0.979	0.403	0.461
kor	0.364	-0.385	0.263	-0.603
lux	0.409	0.518	0.384	0.750
nld	0.364	-0.876	0.339	0.540
nor	0.364	0.986	0.365	0.590
prt	0.358	-0.657	0.368	0.645
swe	0.378	-0.977	0.295	0.572

"*" denote t statistic less than 1.64

Table 6: Statistics for the Exchange Rates

	$logs_t$					$logq_t$				
	$DFGLS$	MZ_α^{GLS}	$\widehat{\alpha}$	$J_0^{.5}$	$J_0^{.2}$	DFGLS	MZ_α^{GLS}	$\widehat{\alpha}$	$J_0^{.5}$	$J_0^{.2}$
aus	-1.334	-4.082	0.956	17	37	-1.981	-6.947	0.923	10(3.0)	21 (7.0)
aut	-2.147	-9.244	0.927	12	20	-2.035	-8.411	0.933	12(3.4)	21 (6.9)
can	-2.068	-11.466	0.955	18	24	-2.305	-17.598	0.945	16(3.2)	19 (4.6)
che	-2.217	-10.585	0.924	11	19	-2.312	-10.605	0.919	10(2.8)	18 (5.7)
deu	-2.191	-10.298	0.928	12	20	-2.094	-8.532	0.935	13(3.8)	22 (7.3)
dnk	-1.802	-6.581	0.955	18	31	-2.046	-8.232	0.940	14(4.2)	23 (7.9)
fra	-1.931	-6.919	0.955	18	29	-2.352	-10.342	0.928	12(3.1)	18 (6.0)
gbr	-1.679	-5.829	0.940	13	28	-1.991	-7.619	0.922	10(2.8)	21 (6.5)
grc	-1.077	-2.589	0.972	26	58	-1.201	-3.234	0.965	21(6.1)	46 (13.0)
ire	-1.593	-5.682	0.961	20	35	-2.199	-8.885	0.903	8(2.2)	17 (5.1)
ita	-1.609	-5.925	0.964	22	37	-2.383	-12.653	0.918	10(2.7)	17 (5.3)
jpn	-2.511	-13.076	0.914	10	16	-2.578	-15.633	0.906	9(2.2)	14 (4.3)
kor	-1.490	-5.735	0.974	29	45	-1.119	-5.723	0.988	5615.2)	79 (20.7)
lux	-1.830	-7.018	0.953	17	29	-1.850	-7.039	0.948	16(4.5)	27 (8.8)
nld	-2.158	-10.083	0.931	12	20	-2.033	-8.843	0.938	13(3.8)	23 (7.4)
nor	-2.021	-7.995	0.936	13	23	-2.223	-9.384	0.922	11(3.9)	19 (8.6)
prt	-0.872	-2.665	0.984	47	83	-1.843	-8.240	0.949	17(5.0)	25 (10.0)
swe	-2.019	-8.485	0.940	14	24	-1.959	-8.082	0.939	14(3.8)	24 (7.7)

Table 5: Output and Inflation in the U.S.

Case	$\widehat{\alpha}_{ML}$	$\widehat{\theta}_{ML}$	DFGLS	DF	MZ_α^{GLS}	Z_α	$\widehat{\alpha}^{GLS}$	$\widehat{\alpha}^{OLS}$	$kmic$	$kbic$	$J_0^{.5}$	$J_1^{.5}$	$J_0^{.2}$	$J_1^{.2}$
gdp	0.954	0.221	-1.576	-2.094	-5.319	-8.020	0.980	0.966	2	1	37	34	63	78
gdpdef	0.950	-0.298	-1.242	-2.834	-3.096	-15.388	0.967	0.895	2	0	14	20	56	48
pcectpi	0.947	-0.283	-1.199	-2.871	-3.089	-16.014	0.962	0.890	9	0	9	18	46	41

Regressions for log(GDP) include a constant and a trend. Only a constant is included in regressions for inflation.