Purchasing Power Parity analyzed through a continuous-time version of the ESTAR model

João Nicolau [*],[1]

School of Economics and Management (ISEG)/Technical University of Lisbon (UTL)/CEMAPRE, Portugal

ARTICLE INFO

ABSTRACT

From the discrete-time Exponential Smooth Autoregressive model, we obtain a continuous-time version that provides new tools for analyzing the Purchasing Power Parity hypothesis.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

One way to test the Purchasing Power Parity (PPP) is by analyzing whether real exchange rates (RER) follow a mean-reverting process. Most estimated models have, in fact, confirmed the existence of an equilibrium level of the RER that is "consistent with a theoretical literature on transaction costs in international arbitrage" (Taylor et al., 2001).

To our best knowledge, all models analyze PPP assuming that the RER are a discrete-time process. The most popular model is the (discrete-time) Exponential Smooth Autoregressive model (ESTAR) proposed by Granger and Teräsvirta (1993). The ESTAR model is used, for example, in Taylor et al. (2001), Kilian and Taylor (2003), Paya and Peel (2006), among others authors.

In this paper, we analyze the conditions under which the discrete-time ESTAR converges in distribution to a diffusion process as the length of the discrete-time intervals between observations goes to zero. We obtain a continuous-time version of the ESTAR process that provides further insights into the mechanism of reversion and into the limit properties of the process. In particular, we conclude that the random walk behavior in the middle of the sample space (the so-called band of inaction) may induce a flat shape in the center of the stationary distribution. This may justify

why the kurtosis of the unconditional empirical distribution of (log-) RER cannot in general be very high (say higher than 3). We also discuss the expected time to leave certain sets, which may be relevant to study the stability of RER and the tendency for RER to maintain in the vicinity of equilibrium value.

The stationary distribution and expected times to leave certain sets are only two examples of functionals that can be easily obtained in diffusion processes but are extremely difficult or even impossible to obtain in a discrete-time setting (especially if the process is non-linear). From a more conceptual point of view, it can be argued that continuous-time formulation is closer to the way that the data are actually generated. As Bergstrom (1993) says: "?…the economy does not move in regular discrete jumps corresponding to the observations—it is adjusting in between observations and it can change at any point of time…"?. This is especially true for RER.

An important advantage of the continuous-time version of the ESTAR model is that the estimated parameters are independent of the frequency of the data. That is, if one uses consistent estimators (such as the ones provided in this paper) and a large time span, the estimates of the model are asymptotically equivalent, whether we consider (say) weekly, monthly, or annual data. This does not hold for a discrete-time model as the results of estimation depend on the frequency of the data. As such, temporal aggregation in discrete-time models from high-frequency data may induce severe bias in the real exchange rate analysis. In fact, Sarno (2000) shows that the nonlinearity present in real exchanges rates is significantly altered by systematic sampling and concludes that "the frequency of the data is crucial for detecting nonlinear mean reversion in RER." Taylor

 * Tel.: +351 213925875.
   E-mail address: nicolau@iseg.utl.pt.
 [1] Postal address: ISEG, Rua do Quelhas 6, 1200-781 Lisbon, Portugal.

(2001) shows that the half-life of deviations suffers from severe bias in the presence of temporal aggregation and that the bias increases as the degree of temporal aggregation increases.

2. The continuous-time version of the ESTAR model

Most studies suggest a nonlinear adjustment of the (log-) RER, say y, according to the ESTAR model

$$y_t = \mu + b_1(y_{t-1}-\mu) + b_2(y_{t-1}-\mu)F(y_{t-1}) + s\varepsilon_t, \tag{1}$$
$$F(x) = 1 - e^{-\theta^2(x-\mu)^2}.$$

Under some mild regularity conditions, it is possible to obtain a continuous-time version of this process, by analyzing the limit process of the stochastic difference Eq. (1), as the length of the discrete-time intervals between observations goes to zero. Let $i\delta$ be the instances at which the process is observed, $(0 \le \delta \le 2\delta \le ...)$ and δ is the interval between observations. When δ changes (as $\delta \to 0$) the parameters b_1, b_2, s, etc. of Eq. (1) should change accordingly. To reflect this fact, we rewrite Eq. (1) as follows:

$$y_{\delta i}-y_{(i-1)\delta} = \left(\frac{b_{\delta 1}-1}{\delta} + \frac{b_{\delta 2}F\left(y_{(i-1)\delta}\right)}{\delta}\right)\left(y_{(i-1)\delta}-\mu\right)\delta + \frac{s_\delta}{\sqrt{\delta}}\sqrt{\delta}\varepsilon_{\delta i} \tag{2}$$

(When $\delta \equiv 1$, we settle $b_{\delta 1} = b_1$, $b_{\delta 2} = b_2$, $s_\delta = s$ etc. and, in this case, Eqs. (1) and (2) are equivalent).

Theorem 1. Assume that (i) the non-stochastic sequences $\left\{\frac{b_{\delta 1}-1}{\delta}\right\}$, $\left\{\frac{b_{\delta 2}}{\delta}\right\}$, $\left\{\frac{s_\delta}{\sqrt{\delta}}\right\}$ converge as follows:

$$\frac{b_{\delta 1}-1}{\delta} \to \phi_1, \ \frac{b_{\delta 2}}{\delta} \to \phi_2, \ \frac{s_\delta}{\sqrt{\delta}} \to \sigma$$

when $\delta \to 0$; (ii) the (stochastic) sequences $\{\varepsilon_{\delta i}\}$ is i.i.d., $E[\varepsilon_{\delta i}] = E[\varepsilon_{\delta i}^3] = 0$, $E[\varepsilon_{\delta i}^4] < \infty$ for all i; (iii) (equality of the initial conditions) $y_0 = X_0$; (iv) $E[(y_\delta - x_1)^4 | y_0 = x_1] < \infty$. In these circumstances, the

process $y_t^\delta = y_{i\delta}$ if $i\delta \le t < (i + 1)\delta$, converges in distribution to X_t when $\delta \to 0$, where X_t is the solution of the stochastic differential equation

$$dX_t = (\phi_1 + \phi_2 F(X_t))(X_t-\mu)dt + \sigma dW_t \tag{3}$$

where $\{W_t\}$ is a Wiener process and $F(x) = 1 - e^{-\theta^2(x-\mu)^2}$.

The proof is in the Appendix. A typical representation of the drift coefficient $a(x) = (\phi_1 + \phi_2 F(x))(x-\mu)$ is presented in Fig. 1, panel B.

Theorem 2. If $\phi_1 + \phi_2 < 0$ then the solution of the stochastic differential Eq. (3) is ergodic and has a stationary density of the form

$$\sigma^{-2}\exp\left\{\frac{(\phi_1 + \phi_2)(x-\mu)^2}{\sigma^2} + \frac{e^{-\theta^2(x-\mu)^2}\phi_2}{\sigma^2\theta^2}\right\}. \tag{4}$$

Furthermore, the solution X is ρ mixing and α mixing.

To prove this theorem, it suffices to apply the lemma given in the Appendix. Notice that, under the condition $\phi_1 + \phi_2 < 0$, the drift coefficient verifies the following crucial condition, defined in the lemma: there exists a $M > 0$ and a $k > 0$ such that $x < -M \Rightarrow a(x) > k$ and $x > M \Rightarrow a(x) < -k$. This condition is the basis for the mean reversion mechanism. It establishes that when the level of X is high (low), say above the constant M (below the constant $-M$), the drift is negative (positive) and so the probability that X decreases (increases) will be high. Therefore, if X is too "high" or too "low" (say if X is outside the set $(-M, M)$), there will be reversion effects that attract it again towards the "central set" $(-M, M)$.

It is worth analyzing the expression of the stationary density (4). We can interpret the density as composed of two multiplying terms (to simplify, assume $\sigma = 1$): (a) $\exp\{(\phi_1 + \phi_2)(x-\mu)^2\}$, $\phi_1 + \phi_2 < 0$ and (b) $\exp\{e^{-\theta^2(x-\mu)^2}\phi_2\theta^{-2}\}$. The first term is related to the normal distribution and has the usual bell-shaped form centered at μ. When $\phi_2 < 0$ (the most natural condition), the second term has a form of an inverted bell (i.e., an upside down bell) also centered at μ. The multiplication of these two terms gives a distribution that is flat at the center of the sample space (an example is given in Fig. 1, panel D). This unusual form has an obvious interpretation: in the middle of the sample space (the so-called band of inaction), where the drift $a(x)$ is approximately zero, the process behaves like a
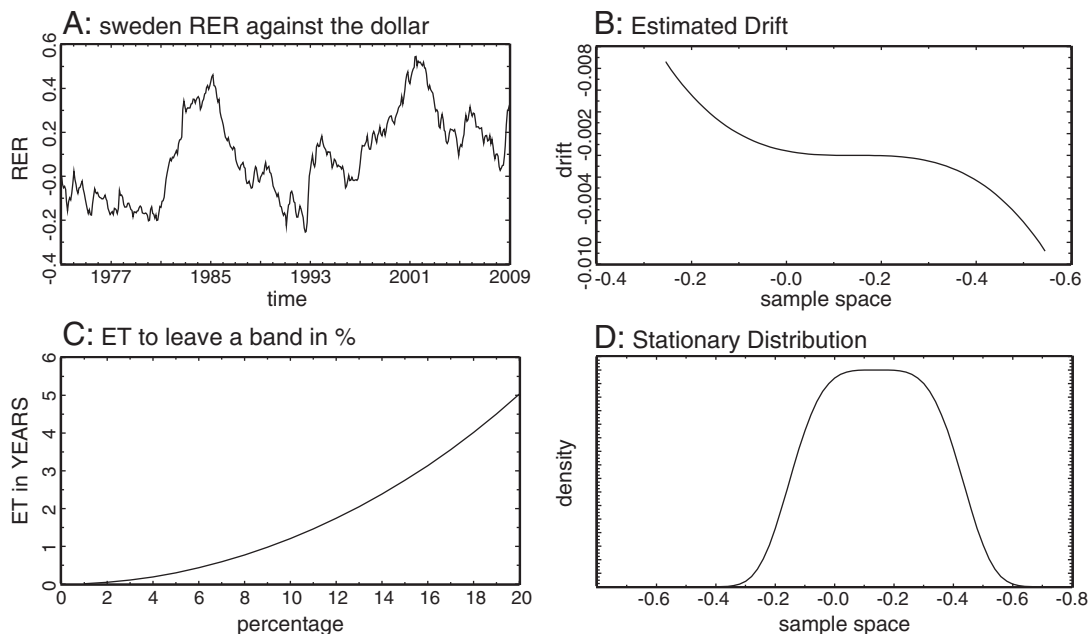

Fig. 1. Sweden RER against the dollar and results of the empirical illustration.

random walk; therefore, inside this interval, there are no sub-intervals more likely than others so the distribution in levels must be flat at the center. It is worth noting that while the distribution of first differences is usually leptokurtic, the distribution in levels may be completely different. In fact, it can be proved that the distribution of the ESTAR process is platykurtic. The conclusion is that, in general, RER cannot exhibit high values of kurtosis, since the random walk behavior in the middle of the sample space induces a flat shape in the center of the distribution. This claim is confirmed in the empirical analysis below.

3. Hitting times and the expected time to leave certain intervals

Suppose that the process is at value x at time zero, i.e., $X_0 = x$. We may ask what the expected time is for the process to reach any element in an arbitrary set, or, which is equivalent, to leave a set, say A that contains x. For example, we may consider the case where x is μ (a central measure of the process) and A is a band of (say) 5 or 10% around μ. Therefore, we may describe how much time (in mean) the process closely stays around a central measure of the process. This analysis provides further insights about the stability of the RER and how much time the RER stays in the neighborhood of the PPP equilibrium.

Analytically, the issue is to find $E(T|X_0 = x)$ where $T = \inf\{t \geq 0 : X_t \notin A\}$ and $x \in A$. It can be shown that the functional $w(x) = E(T|X_0 = x)$ satisfies the partial differential equation $Lw(x) = -1$, where $L = a(x)\frac{\partial}{\partial x} + \frac{1}{2}\sigma^2\frac{\partial^2}{\partial x^2}$, with the boundary conditions $w(l) = w(u) = 0$. The solution of this problem is given in Karlin and Taylor (1981):

$$E(T|X_0 = x) = 2\Big\{k(x)\int_x^u (S(u)-S(\xi))m(\xi)d\xi$$
$$+ (1-k(x))\int_l^x (S(\xi)-S(l))m(\xi)d\xi\Big\}$$

where $k(x) = (S(x)-S(l))/(S(u)-S(l))$, $S(x) = \int^x s(u)du$, $s(x) = \exp\{-\int^x 2a(u)\sigma^{-2}du\}$, and $m(x) = (s(x)\sigma^2)^{-1}$. The expressions $s(x)$ and $m(x)$ are known in closed form, while $S(x)$ must be calculated numerically.

4. An empirical illustration

To illustrate the model, we estimate the Sweden RER against the dollar using monthly data from January 1973 to May 2009 (437 observations). The data were first normalized on the beginning of the sample and transformed in logarithm. The time series is therefore $X_t = ln(S_t)$, where S_t is the RER (notice that at the beginning of the sample, we have $X_1 = 0$). The data are available at site http://pascal.iseg.utl.pt/nicolau/sweden.xls, and the source is the International Financial Statistics of the International Monetary Fund and Financial Statistics of the Federal Reserve Board.

The transition (or conditional) densities of X required to construct the exact likelihood function are unknown. To estimate the parameters of Eq. (3), we considered the simulated maximum likelihood estimator suggested in Nicolau (2002) (with $N = 100$ and $S = 50$). The method is consistent and fully efficient as $N, S \rightarrow \infty$. We obtained the following results (the parameters have monthly interpretation, that is, we suppose that the X process is observed at instants $\{i\Delta, i = 0, 1,..., n\}$, with a step of discretization of $\Delta = 1$).

We impose the restriction $\phi_1 = 0$, since ϕ_1 was not found to be significantly different from zero. From Table 1, we obtained the following estimates: (1) $\hat{a}(x) = \hat{\phi}_2 F(x)(x-\hat{\mu})$ (see Fig. 1, panel B); (2) expected time in years to leave a band of p%, give that at initial time the process is at state $\hat{\mu}$ (panel C); and (3) stationary distribution built from Eq. (3) (where the normalizing constant of (4), unknown in closed form, was found by numeric integration) (panel D).

Table 1
Estimation results.

	ϕ_1	ϕ_2	μ	θ	σ
Estimates (standard errors)	0 (NA)	−1.33 (0.64)	0.143 (0.048)	0.32 (0.09)	0.0262 (0.001)

The estimated drift is flat and approximately zero in the center of the sample space, positive when X is low and negative when X is high. This shape of the drift coefficient suggests the existence of an equilibrium level at $\hat{\mu} = 0.143$, in the vicinity of which the RER behave like a random walk, becoming increasingly mean reverting as RER deviates from equilibrium.

With regard to the distribution (Fig. 1, panel D), see our discussion in Section 2. It is interesting to observe that the kurtosis estimate for the Sweden RER is 2.11. Finally, in panel C, we present $E(T_{p \times 100}|X_0 = \hat{\mu})$ as a function of $p \times 100$ where

$$T_{p \times 100} = \inf\{t \geq 0 : X_t \notin (\mu-p, \mu+p)\}.$$

One reads, for example, that $E(T_{10}|X_0 = \hat{\mu})$ is approximately 1 year, that is, the Sweden RER will take about 1 year (on average) to leave the set $(\mu-0.10, \mu+0.10)$ (i.e., a band of ±10%) given that the process at initial time is at state $\hat{\mu}$ (equilibrium value).

Acknowledgments

Appendix A

Proof of Theorem 1. We apply theorem 2.2 of Nelson (1990), which is based on four assumptions A2–A5. The A2–A4 assumptions are trivially satisfied (the initial conditions are the same and Eq. (3) satisfies the Lipschitz and the linear growth bound conditions—a proof of this claim can be given upon request). The crucial assumption A5 still has to be verified. Under the hypotheses of our theorem, we have three results: (A), (B), and (C) (see below). These results imply, by definition, A5. We have after simplifications (noting that $E(\varepsilon_\delta) = E(\varepsilon_\delta^3) = 0$):

(A)

$$\lim_{\delta \to 0} \sup_{x<R} \frac{E\left[(y_\delta-y_0)^4|y_0 = x\right]}{\delta}$$
$$= \lim_{\delta \to 0} \sup_{x<R}\Bigg\{\left(\frac{b_{\delta1}-1}{\delta} + \frac{b_{\delta2}F(x)}{\delta}\right)^4(x-\mu)^4\delta^3$$
$$+ 6\left(\frac{b_{\delta1}-1}{\delta} + \frac{b_{\delta2}F(x)}{\delta}\right)^2(x-\mu)^2\delta\left(\frac{s_\delta}{\sqrt{\delta}}\sqrt{\delta}\right)^2 E\left(\varepsilon_\delta^2\right)$$
$$+ \frac{1}{\delta}\left(\frac{s_\delta}{\sqrt{\delta}}\sqrt{\delta}\right)^4 E\left(\varepsilon_\delta^4\right)\Bigg\} = 0;$$

Let $a_\delta(x)$ and $a(x)$ be the drifts of $y_{i\delta}$ and X_t, respectively. Then

B)

$$\lim_{\delta \to 0} \sup_{|x|<R} |a_\delta(x)-a(x)|$$
$$= \lim_{\delta \to 0} \sup_{|x|<R}\left|\left(\frac{b_{\delta1}-1}{\delta} + \frac{b_{\delta2}F(x)}{\delta}\right)(x-\mu)-(\phi_1+\phi_2 F(x))(x-\mu)\right|$$
$$= \lim_{\delta \to 0} \sup_{|x|<R}\left|\left(\frac{b_{\delta1}-1}{\delta}-\phi_1 + \frac{b_{\delta2}F(x)}{\delta}-\phi_2 F(x)\right)(x-\mu)\right| = 0;$$

Let $\left(s_\delta/\sqrt{\delta}\right)^2$ and σ^2 be the "diffusion" coefficients of $y_{i\delta}$ and X_t, respectively. Thus,

C)

$$\lim_{\delta \to 0} \sup_{|x|<R} \left|\left(s_\delta/\sqrt{\delta}\right)^2 - \sigma^2\right| = 0. \qquad \blacksquare$$

Lemma. *Let $X = \{X_t, t \geq 0\}$ be a diffusion process, with state space $I = (-\infty, \infty)$, $l = -\infty$, $r = \infty$, governed by the Itô stochastic differential equation $dX_t = a(X_t)dt + \sigma dW_t$, $(X_0 = x, \sigma > 0)$, where $\{W_t, t \geq 0\}$ is a (standard) Wiener process, x is a random value F_0-mensurable independent of W_t and a is continuously differentiable. If there exists a $M > 0$ and a $k > 0$ such that*

$$x < -M \Rightarrow a(x) > k \text{ and } x > M \Rightarrow a(x) < -k,$$

then X is ergodic and possesses a stationary distribution proportional to $\{\int^x 2a(u)/\sigma^2 du\}$. Furthermore, X is ρ mixing and α mixing.

Proof. Let $s(x) = \exp\{-\int_{z_0}^x 2a(u)/\sigma^2 du\}$ be the scale density function (z_0 is an arbitrary point inside I) and $m(x) = (\sigma^2 s(x))^{-1}$ the speed density function. Let $S(-\infty, x] = \lim_{x_1 \to -\infty} \int_{x_1}^x s(u)du$ and $S[x, \infty) = \lim_{x_2 \to \infty} \int_x^{x_2} s(u)du$ where $-\infty < x_1 < x < x_2 < \infty$. Under the conditions of the lemma, it is not difficult to conclude that (R1) $S(-\infty, x] = S[x, \infty) = \infty$ for $x \in I$; (R2) $\int_{-\infty}^{\infty} m(x)dx < \infty$; (R3) $\lim_{x \to \infty} \sup a(x)/\sigma < 0$, $\lim_{x \to -\infty} \sup a(x)/\sigma > 0$. (R1) and (R2) imply that X is ergodic and the invariant distribution P^0 has density $m(x)/\int_{-\infty}^{\infty} m(u)du$ with respect to the Lebesgue measure (Skorokhod, 1989, theorem 16). (R3) implies that the process is ρ mixing and α mixing (Chen et al., 2009). $\qquad \blacksquare$

References

Bergstrom, A., 1993. The ET interview: A.R. Bergstrom. In: Phillips, P. (Ed.), Models, Methods and Applications of Econometrics. B. Blackwell, Cambridge, Mass.

Chen, X., L. Hansen and M. Carrasco, 2009, Nonlinearity and Temporal Dependence, Working Paper 2009s–17 Scientific Series, Cirano.

Granger, Teräsvirta, 1993. Modelling Nonlinear Economic Relationships. Oxford University Press, Oxford.

Karlin, S., Taylor, H., 1981. A Second Course in Stochastic Processes. Academic Press.

Kilian, L., Taylor, M., 2003. Why is it so difficult to beat the random walk forecast of exchange rates? Journal of International Economics 60 (1), 85–107.

Nelson, D., 1990. Arch models as diffusion approximations. Journal of Econometrics 45, 7–38.

Nicolau, J., 2002. New technique for simulating the likelihood of stochastic differential equations. The Econometrics Journal 5 (1), 91–103.

Paya, I., Peel, D., 2006. On the speed of adjustment in ESTAR models when allowance is made for bias in estimation. Economics Letters 90 (2), 272–277.

Sarno, L., 2000. Systematic sampling and real exchange rates. Review of World Economics 136 (1), 24–57.

Skorokhod, A., 1989. Asymptotic methods in the theory of stochastic differential equation. Translation of Mathematical Monographs, 78. American Mathematical Society.

Taylor, A., 2001. Potential pitfalls for the purchasing-power-parity puzzle? Sampling and specification biases in mean-reversion tests of the law of one price. Econometrica 69 (2), 473–498.

Taylor, M., Peel, D., Sarno, L., 2001. Nonlinear mean-reversion in real exchange rates: towards a solution to the purchasing power parity puzzles. International Economic Review 42 (4), 1015–1042.