Lancaster University Management School
Working Paper
2005/016

# Temporal aggregation of an ESTAR process: some implications for purchasing power parity adjustment

David Peel and Ivan Paya

The Department of Economics
Lancaster University Management School
Lancaster LA1 4YX
UK



The LUMS Working Papers series can be accessed at http://www.lums.lancs.ac.uk/publications/
LUMS home page: http://www.lums.lancs.ac.uk/

# Temporal Aggregation of an ESTAR Process: Some Implications for Purchasing Power Parity Adjustment

Ivan Paya[a,1], David A. Peel[b]
[a]Departamento Fundamento Analisis Economico, University of Alicante, 03080 Alicante, Spain
[a]Lancaster University Management School, Lancaster, LA1 4YX, UK, United Kingdom.

**Abstract**

Nonlinear models of deviations from PPP have recently provided an important, theoretically well motivated, contribution to the PPP puzzle. Most of these studies use temporally aggregated data to empirically estimate the nonlinear models. As noted by Taylor (2001), if the true DGP is nonlinear, the temporally aggregated data could exhibit misleading properties regarding the adjustment speeds. We examine the effects of different levels of temporal aggregation on estimates of ESTAR models of real exchange rates.

Keywords: ESTAR, Real Exchange Rate, Purchasing Power Parity, Aggregation.
JEL classification: F31, C22, C51

[1]**Corresponding author.** E-mail address: ivanpaya@merlin.fae.ua.es Tel.: +34 965903614. Fax: +34 965903898.

# 1 Introduction

Recently, a number of authors have reported empirical results that show that after allowing for nonlinearities there is apparent mean reversion in real exchange rates. The nonlinear models reported have been estimated on data sampled at different levels of aggregation, namely monthly, quarterly and annual (see, e.g., Michael et al., 1997; Baum et al., 2001; Taylor et al., 2001; Kilian and Taylor, 2003; and Paya et al., 2003). As noted by Taylor (2001), much of the data employed in empirical work is temporally aggregated.

There are some interesting issues raised by this work. A natural concern is that the estimated nonlinear models may exhibit misleading properties when the underlying data generating process operates at a higher frequency than the observed data. One worry is that temporal aggregation may imply the disappearance of nonlinearity. Another concern is that after temporal aggregation the measured adjustment speeds based on nonlinear model estimates may be biased.

The current paper addresses these concerns within the context of a specific form of nonlinear model, namely the Exponential Smooth Autoregressive (ESTAR) model, that has been widely used to model real exchange rates. We follow a set up similar to Taylor (2001) but that differs in one important respect. We generate artificial data at the, possibly unobservable, high frequency from an ESTAR model and temporally aggregate these data to frequencies of interest in applied work. We then fit ESTAR models to the temporally aggregated data at hypothetical monthly, quarterly and annual frequencies. This differs from Taylor who estimates linear models on the temporally aggregated data and shows that the linear estimates of adjustment speeds can be substantially downward biased.

Based on Monte Carlo simulation we show that ESTAR type nonlinearites are usually preserved under the temporal aggregation schemes we consider.[2] However the dynamic structure of the best fitting models changes. In fact the best fitting models in our simulations, for monthly, quarterly or annual frequency, tend to take the form researchers have found to fit well on actual data of the same frequency. This fact provides evidence in favor of temporal aggregation and complements the direct evidence referred to in Taylor (2001) in his discussion of the IMF's data compilation. Furthermore comparison of the measured speed of response to shocks with models estimated on the temporally aggregated data and the true DGP shows that the measured speed of adjustment declines the more aggregated the data.

The rest of the paper is organized as follows. In section 2 we set out the DGP for highest frequency data, our Monte Carlo methodology, the linearity tests and the effect of temporal aggregation on nonlinear estimates of an ESTAR model. Section 3 compares the Monte Carlo results with actual estimates.

[2] Very little work has been done on the effects of aggregation on non-linear time series models. Granger (1991), and Granger and Lee (1999) are notable exceptions. However, their analysis does not consider nonlinear processes involving symmetric adjustment in models that can exhibit near unit root behaviour.

In section 4 we examine, employing nonlinear impulse response functions, the speeds of adjustment to shocks obtained in the DGP and the estimated temporally aggregated ESTAR models. Finally, section 5 summarizes our main conclusions.

# 2 The true structural model and the effect of time aggregation on estimated nonlinear parameters

We assume that at the highest frequency the DGP is given by an ESTAR model of Ozaki (1985). A smooth rather than discrete adjustment mechanism is chosen for two reasons. First a smooth adjustment process is suggested by the theoretical analysis of Dumas (1992). Second, as postulated by Terasvirta (1994) and demonstrated theoretically by Berka (2002), in aggregate data, regime changes may be smooth rather than discrete given that heterogeneous agents do not act simultaneously even if they make dichotomous decisions. We assume that the ESTAR model which describes the DGP for modelling PPP deviations at the highest data frequency has the simplest possible lag structure within the class of ESTAR models and is given by

$$y_t = e^{-\gamma y_{t-1}^2} y_{t-1} + u_t \tag{1}$$

where $\gamma$ is a positive constant and $u_t$ is a white noise disturbance term with standard error ($se$).

In fact the DGP given by equation (1) is that typically reported in empirical studies of monthly data, the highest frequency observable in practice.

Figure 1 is a deterministic plot of the relationship between $\Delta y = y_t - y_{t-1}$ and $y_{t-1}$ obtained from (1). We observe in Figure 1 that for small deviations from equilibrium, adjustment may be modelled as a unit root process - "the optimality of doing nothing" - but for large deviations from equilibrium there is mean reversion. If the process spends a significant proportion of time in or near the unit root region, it will exhibit strong persistence and near unit root behavior.

We simulate data from the ESTAR model (1) where the disturbance term, $u_t$, is assumed to be normally distributed.[3]

Following Taylor (2001) we create arithmetic temporal aggregates from the simulated data as[4]

[3] We also consider nonnormal disturbances such as t-Student with 18 degrees of freedom that in previous research appears to match the nonnormality of residuals (see Paya and Peel 2003). Results were qualitatively unchanged.

[4] If the data is in logarithmic form, then $y_t^*$ is the geometric mean instead of the arithmetic mean of the real exchange rates. We compared the correlation between the arithmetic and geometric means conditional on some price processes. The correlations were close to unity and the results qualitatively similar. Given this for simplicity we follow Taylor (2001) and employ the arithmetic mean for the temporally aggregated data.

$$y_t^* = \frac{(y_t + y_{t-1} + y_{t-2} + ...... + y_{t-(i-1)})}{i} \qquad (2)$$

where $i = 2, 3, 12$.

Two different assumptions about the true DGP are made. First, we assume that the true DGP is a nonlinear 'monthly' ESTAR process and simulate from this 120,000 observations. We replicate this experiment 1,000 times.

The range of standard deviations of the disturbance term is calibrated on the monthly estimates of equation (1), the highest aggregate data frequency available to researchers (see e.g., Taylor et al., 2001; and Venetis et al., 2002). These studies report standard errors of around 0.035. For purposes of comparison we also simulate series with a much lower standard deviation than found in the monthly data and employ values of 0.01 and 0.035. The adjustment parameter is given the values of $\gamma = 0.5, 1$. The estimates obtained in actual monthly data tend to fall in this range.

Aggregating these observations three times, $i = 3$ (quarterly), or twelve times, $i = 12$ (annual), yields 1,000 samples of 40,000 and 10,000 observations, respectively. These samples will be used to analyze the 'large sample' behaviour of aggregated nonlinear 'monthly' ESTAR models. To analyze the small sample properties, we employ the same method but limit the sample sizes to 120 for 'quarterly' ($i = 3$) aggregation and 200 for 'annual' ($i = 12$) as these span the most common used samples in the literature.[5]

The second assumption made is that the true DGP given by (1) is for data generated at either a fortnightly or ten days frequency that is aggregated to monthly data ($i = 2$, or 3) respectively. Again 120,000 observations are simulated from (1) 1,000 times. In this case the standard errors of the residuals in the true DGP are chosen as $\sigma = 0.024$, 0.028 so that the standard errors of the residuals in the temporally aggregated data, $i = 2, i = 3$ match those found in actual monthly estimates. Values of $\gamma = 0.3, 0.4$ were employed which produced values of the speed of adjustment parameter in the aggregate data similar to those observed in empirical work. In this exercise, the large sample analysis was done with 10,000 observations of the aggregated data[6] and the small sample analysis with 360 observations matching the sample size of monthly data on real exchange rates available from the post Bretton Woods period and around the length of sample that has typically been employed in previous empirical analysis.

[5] Samples of real exchange rates of 120 at quarterly data are available for the post Bretton-Woods period. At annual frequency the longest data set available is from 1792 in the case of Dollar/Pound and Dollar/French Franc (see Lothian and Taylor 1996).

[6] The results employing 60,000 or 40,000 appeared essentially the same than on a sample using 10,000. We report results on samples of 10,000 as it was computationally much less time consuming.

## 2.1 Testing for nonlinearity

Recent research has developed new testing procedures for the null hypothesis of a unit root process against the alternative hypothesis of a nonlinear exponential smooth transition autoregressive (ESTAR) process, which is globally mean reverting. Kapetanios et al. (2003), (KSS hereafter), derived a unit root test against a nonlinear (and asymptotically stationary) alternative.[7] This test has better power than the standard Dickey-Fuller test in the region of the null. They test the null hypothesis of a linear model, $H_0 : \gamma = 0$.

KSS (table 3) report the power of their test for different parameter values in the case where the residual term follows a standard normal distribution. They show that the power of the test depends upon the values of the parameters in the ESTAR form. The standard deviation of the error term in our simulated processes is 0.035 and 0.01, and the values of $\gamma = \{1, 0.5\}$.

Kiliç (2003) developed an alternative testing method to detect the presence of nonstationarity against nonlinear but globally stationary STAR process that differs from KSS in the way it deals with the nuisance parameter that occurs under the null. As the author claims, the advantage of Kiliç procedure over KSS is twofold. First, it computes the test statistic even when the threshold parameter needs to be estimated in addition to the transition parameter. Second, it claims to have higher power.

Table 1 reports the power of the KSS and Kiliç tests for large and small sample sizes.[8] The power of the tests are low for small sample sizes so that the results of applying the KSS test to our processes should be interpreted with this caveat in mind.

We apply both the KSS[9] and Kilic[10] tests to our aggregated nonlinear ESTAR processes. Table 2 presents the results and displays the proportion of times that each individual test as well as both tests would reject the null hypothesis of unit root against the alternative of an STAR. For large samples both tests would always reject the null. With regard to small samples, the higher the standard error ($se$) and the higher the speed of adjustment ($\gamma$), the higher is the proportion of rejections. It is also worth pointing out that the greater the

[7] KSS examine the properties of their test under three different assumptions of stochastic processes with nonzero mean and/or linear deterministic trend. In the cases where $y_t^*$ exhibits significant constant or trend, $y_t^*$ should be viewed as the de-meaned and/or de-trended variable.

[8] In order to examine the power of the KSS test, we simulate model (18) in KSS. Please note that KSS notation differs from ours. They use the parameter $\gamma$ to denote the autoregressive process of the dependent variable and the parameter $\theta$ for the speed of adjustment. To exactly match our parameter values KSS simulations should be those with $\phi = 0, \gamma = -1, \theta = \{0.01^2, 0.5x0.01^2, 0.035^2, 0.5x0.035^2\}$

[9] In order to apply the KSS to the aggregated process we first regress $y^*$ on a constant and trend. In cases where the constant and/or trend were significant we demeaned or detrended the series and use the appropriate critical values.

[10] Kiliç suggests that making the interval too wide could make the transition function to be flat for large values of $\gamma$. We have then decided to use an interval for $\gamma$ according to values usually found in our simulation results for each degree of aggregation. The values of $C$ have been selected as the corresponding to the ordered values of $|z|$ and discard 10% of the highest and smallest values.

degree of temporal aggregation (e.g., monthly to annual as apposed to monthly to quarterly) the greater the proportion of rejections. However, we must interpret these results with some caution as the power of the tests imply that non-rejection of the null might often occur when the process is ESTAR.

In the next section we will examine the properties of the nonlinear ESTAR estimation. One approach would be to first test for ESTAR nonlinearity using both the KSS and Kilic tests and if they do not reject the unit root hypothesis then do not carry any further nonlinear estimation. However, following our previous analysis on the power of these tests, a researcher could carry on with the estimation process even though rejection of the unit root hypothesis might not occur. We therefore analyze the effects that temporal aggregation has on the ESTAR estimation process as a separate exercise.

## 2.2 Nonlinear ESTAR Estimation

On the aggregated data we estimate by nonlinear least squares the following model

$$y_t^* = a + B(L)y_{t-1}^* e^{-\gamma(y_{t-1}^* - a)^2} + v_t \quad (3)$$

where B(L) is a polynomial lag operator of order up to five which rendered the disturbance term $v_t$ empirical white noise,[11] and $a$ is a constant. Empirical marginal significance levels of the estimated parameter $\gamma$ are obtained through Monte Carlo simulation as it is not defined under the null. In particular, the model is assumed to follow a unit root linear autoregressive process and then a nonlinear ESTAR specification (equation 3) is estimated, computing the appropriate confidence interval of significance for $\gamma$.

First, we examine the results obtained in the case of the large samples described above. We observe in the results reported in Tables 3a, 3b, 3c and 3d that time aggregation induces higher order autoregressive terms in the fitted models at lower frequencies than occur in the DGP. Moreover, the additional autoregressive structure induced by time aggregation seems to have a limiting number of terms. The second order autoregressive term is always significant. Terms in an autoregressive process of order three are significant at least 95% of the time except for i=2 when it falls to 59%. Higher order terms exhibit a steep fall in significance. The significance of the AR(4) parameter varies between 37% and 7% with that of the AR(5) parameter between 5-7%. The order of the autoregressive structure appears to be independent of the range of standard errors of the disturbance term and the speed of adjustment parameters imposed in the true DGP in our simulations.

The regression standard error and the point estimate of the speed of adjustment parameter, $\gamma$, increase with the degree of aggregation. The speed of adjustment parameter is always significant in the large sample estimates. Another feature of the time aggregation is the finding of significant LM test for

[11] On the basis of the LM test of Eitrheim and Terasvirta (1996).

ARCH. The greater the degree of aggregation and the higher the standard error of the disturbance term in the DGP the more accentuated the finding of a significant LM test for ARCH. Noting that the LM test for ARCH is a test for model misspecification and that the errors in the DGP do not exhibit ARCH, this suggests that specification (3) may become less parsimonious as an appropriate way of modelling the temporally aggregated process (1) as the degree of aggregation increases. We also note that the lower the frequency and the higher the standard error of the disturbance term the lower the goodness of fit parameter $R^2$.

When the estimations are undertaken with smaller samples of observations of 120, 200 and 360, corresponding to quarterly, annual and monthly data employed in empirical studies, the nonlinear estimates of (3) show the following features. The fitted ESTAR exhibits significant AR(2) structure between 50 and 89 percent of the time for i=2,3,12 dependent upon the noise and the speed of adjustment in the true DGP. Autoregressive terms of order greater than two are significant less than ten percent of times. Significant LM tests for ARCH are not found in 90% of the fitted models. The estimated speed of adjustment parameters are higher than in the large sample simulations[12] with larger standard errors and approximately forty percent are significant at the 5% significance level. Consequently, in small sample estimates of nonlinear ESTAR models, on temporally aggregated data, we could erroneously reject the hypothesis that the true DGP follows a nonlinear process.[13]

Nonlinear ESTAR models have been reported at various levels of aggregation and the reported empirical results conform with those obtained on the simulated data. Kilian and Taylor (2003) report AR(2) structure in all ESTAR models fitted to quarterly data for seven OECD economies. Michael et al. (1997) report AR(2) structure employing annual data. Also significant LM tests for ARCH are rarely reported.

# 3 Further comparison between simulated data and empirical estimates from actual data

We now proceed to compare further the empirical results obtained from simulated data with those obtained from actual data. Table 4 presents monthly estimates of ESTAR models for seven bilateral real exchange rates against the Dollar in the post Bretton Woods era taken from Venetis et al. (2002). The estimated model corresponds to that of Equation (3). The estimates of $\gamma$ are between 0.16 and 0.8 and the standard deviation of the regressions is around 0.033. We added an additional column, where the $p$-value of the second AR term in the estimates is included. For the majority of the cases, PPP devi-

[12] See Paya and Peel (2004a) for a discussion on the upward bias of the ESTAR estimates in small samples.

[13] Granger and Lee (1999) examine the effects of time aggregation on nonlinearity tests drawing a similar conclusion. Nonlinearity could be rejected when the model has been temporally aggregated.

ations appear parsimoniously described by the simple ESTAR structure given by equation (1). However, it appears that in the case of the Dollar/Yen at the five percent level and the Dollar/Pound and Dollar/Lira at the fifteen percent level, the second AR term plays a significant role. Simulations presented above show that time aggregation induces AR(2) structure in the estimated nonlinear process.

Empirical results at different levels of aggregation ($i = 3, i = 12$) are reported in Tables 5 and 6. The quarterly estimates are taken from Kilian and Taylor (2003). We also present annual estimates of Equation (3) for the Dollar/Pound and the Dollar/Franc for two hundred years derived by Lothian and Taylor (1996) and analyzed by Michael et al. (1997). We note that this data set spans many changes in exchange rate regimes so the results need to be interpreted with that caveat in mind. The Dollar/Deutsche Mark is for the Gold Standard -data source- reported in Paya and Peel (2004b). We observe that the estimates of $\gamma$ are higher than at monthly frequency and similar to those suggested by the simulation exercise above. We also note that the autoregressive structures have a significant AR(2) component.[14] This is interesting given our Monte Carlo showed that in over fifty percent of simulations at "quarterly aggregation" and seventy five percent of simulations at "annual aggregation" gave rise to this specification.

## 4 Generalized impulse response functions

One of the major objections to PPP following a random walk is the counter-intuitive conclusion that shocks persist forever. In addition, even if a stationary linear model could be specified with near unit root behaviour, the percentage absorption of shocks over time will be the same regardless the shock magnitude. In this section, we analyze the persistence properties of temporally aggregated exponential smooth transition models. A number of properties of the impulse response functions of linear models do not carry over to the nonlinear models.[15]

The Generalized Impulse Response Function (GIRF) introduced by Koop, Pesaran and Potter (1996) is defined as the average difference between two realizations of the stochastic process $\{y_{t+h}\}$ which start with identical histories up to time $t-1$ (initial conditions) but one realization is "hit" by a shock at time $t$ while for the other (the benchmark profile) no shock occurs. The GIRF of Koop et al. (1996) is defined as,

[14] In the case of the Dollar/Franc the AR(2) term is insignificant but the residuals exhibit better properties. It is worth noting that these estimations span a long period of time with different exchange rate regimes. Even though those nonlinear estimates have recently been proved to be robust (see Lothian and Taylor, 2004; and Paya and Peel, 2004c) under heteroskedastic residuals they must be taken with caution as there might be some unexplored effects of regime changes on the nonlinear parameters.

[15] In particular, impulse responses produced by nonlinear models are; a) history dependent, so they depend on initial conditions, b) dependent on the size and sign of the current shock, and c) they depend on future shocks as well.

$$GIRF_h(h, \delta, \omega_{t-1}) = E(y_{t+h}|u_t = \delta, \omega_{t-1}) - E(y_{t+h}|u_t = 0, \omega_{t-1}) \quad (4)$$

where $h = 1, 2, ..,$ denotes horizon, $u_t = \delta$ is an arbitrary shock occurring at time $t$ and $\omega_{t-1}$ defines the history set of $y_t$. Given that $\delta$ and $\omega_{t-1}$ are single realizations of random variables, expression (4) is considered to be a random variable.

Note that (4) is general enough to allow multiple interpretations.[16] Here we choose to condition upon "all past histories". Simulation of various shock sizes ($\delta$ values) will then illustrate the possible differences in persistence arising in the ESTAR model from different shock sizes.

Since analytic expressions for the conditional expectations involved in (4) are not available for $h > 1$, we use stochastic simulation (Gallant et al., 1993; and Koop et al., 1996; for a detailed description} to approximate function (4).

Given a particular value of the log real exchange rate at time t, a shock of $k$ percent to the level of the real exchange rate involves augmenting $y_t$ additively by $\ln(1 + k/100)$. Hence, $u_t = \delta$ at time $t$ and we choose $\delta = \ln(1 + k/100)$ with $k = 5, 20, 30$. The particular choice of $\delta$'s allows us to compare and contrast the persistence of large and small shocks. For each history, we construct 5000 replications of the sample paths $\hat{y}_0^*, ..., \hat{y}_h^*$ based on $u_t = \delta$ and $u_t = 0$ by randomly drawn residuals as noise for $h \geq 1$. The difference of these paths is averaged across the 5000 replications and it is stored. At the end, we average across histories.

The persistence of the shocks could be evaluated as suggested by Koop et al. (1996), using the dispersion of the distribution of (4) as horizon $h$ increases.[17] However, the main issue is to compute how many periods ($h$) are necessary for the impulse response function to be "significantly" reduced.

In the case of nonlinear models, monotonicity need not hold.[18] Hence, we calculate the $x - life$ of shocks for $(1 - x) = 0.25$, 0.50 and 0.80 where $1 - x$ corresponds to the fraction of the initial effect $u_t$ that has been absorbed.[19]

We examine the implied speeds of adjustment to shocks for the Monte Carlo experiments in Section 2.[20] Table 7 reports the results of applying this procedure in the case of the large and small samples outlined in section 2. Larger shocks

[16] For example we can condition upon the specific realization of $\Omega_{t-1}$ that reads "all past values" $\omega_{t-1} = \{y_1, y_2, ..., y_{t-1}\}$ and treat $u_t$ as randomly chosen. Or we can condition on "all positive past histories" and $\omega_{t-1} = \{y_i : 1 \leq i \leq t-1 \text{ and } y_i > 0\}$. Accordingly there are histories defined by a fixed band set $-b < y_i < b$ etc. Of course the error shock hitting the model at time $t$ could also belong to some positive, negative or fixed band set just to mention a few possible cases.

[17] See the working paper version of the article (Paya and Peel 2004d) for a detailed discussion of this issue.

[18] We thank an anonymous referee for pointing out this fact. For a full discussion on different measures of half-life shocks and estimating prodedures see Murray and Papell (2002) and Killian and Zha (2002).

[19] See Van Dijk, Franses and Boswijk (2000, p.7)

[20] We refer the reader to the working paper version of the article (Paya and Peel 2004d) for a detailed description of the procedure followed to obtain the GIRF of the temporally aggregated data in the Monte Carlo experiment.

always imply faster adjustments and the reduction in the time needed to absorb fraction $(1-x)$ of different size shocks depends on the proportion $(1-x)$. In other words, if the shock increases from 5% to 20% the reduction in the time needed to absorb 25% of both shock is not the same as the reduction in time needed to absorb 50% of the shocks. Moreover, in either large and small samples the time needed to absorb $(1-x)$ of the shock increases with the aggregation process.

It is also worth pointing out that the upward bias obtain in the small sample estimates discussed in section 2 imply faster adjustment to shocks in the small sample case than in the large sample case.

In order to compare our simulation results with actual estimates, Tables 4, 5 and 6 display the half-life shocks $(1-x=50\%)$ for the nonlinear models estimated on actual data reported in those tables. Employing the simulations results as a benchmark, we can then use these empirical estimates of half-lives of shocks to try to approximate the nature of the true DGP of PPP deviations. We will concentrate on the speed of adjustment to shocks of the Dollar/Pound, Dollar/French Franc, and Dollar/Deutsche Mark. The difference between the speed of adjustment to shocks in the monthly and annual data is around twenty four months for the three different currencies. The difference between the adjustment at quarterly and annual data is either zero or twelve months. This pattern is the one followed by the Monte Carlo results when we aggregate a true DGP from monthly to quarterly and annual data.

# 5 Conclusions

Nonlinear models of deviations from PPP have recently provided an important, theoretically well motivated, contribution to the PPP puzzle. Most of these studies use temporally aggregated data to empirically estimate the nonlinear models. In this paper we have assumed the true DGP at the highest data frequency is an ESTAR model. Given this model we have generated artificial data and examined the effects of different levels of temporal aggregation on estimates of ESTAR models of real exchange rates. Our principal findings are that ESTAR nonlinearities are generally preserved in the temporally aggregated data, though the lag structure changes, and that the implied speed of adjustment to shocks declines the more aggregated the data.

## Acknowledgements

The authors are grateful for helpful and constructive comments from the editor and two anonymous referees who materially improved this paper. Naturally remaining errors are our responsibility. The authors are grateful to participants of the Understanding Evolving Macroeconomy Annual Conference, University College Oxford 15-16 September 2003. I. Paya acknowledges financial support from the Instituto Valenciano de Investigaciones Economicas (IVIE). David Peel gratefully acknowledges financial support by ESRC grant L/138/25/1004.

**REFERENCES**


Baum, C.F., J.T. Barkoulas and M. Caglayan (2001), Nonlinear adjustment to purchasing power parity in the post-Bretton Woods era', *Journal of International Money and Finance,* 20, 379-399.

Berka, M. (2002), 'General equilibrium model of arbitrage, trade and real exchange rate persistence'. *Mimeo. University of British Columbia.*

Dumas, B. (1992), 'Dynamic equilibrium and the real exchange rate in a spatially separated world', *Review of Financial Studies,* 5, 153-80.

Eitrheim, O, and T. Terasvirta (1996), 'Testing the adequacy of smooth transition autoregressive models', *Journal of Econometrics,* 74, 59-75.

Gallant AR, Rossi PE, Tauchen G (1993) Nonlinear Dynamic Structures. Econometrica 61:871 908

Granger, C.W.J. (1991), 'Developments in the nonlinear analysis of economic series', *Scandinavian Journal of Economics*, 93, 263-272.

Granger, C.W.J. and T. Terasvirta (1999), 'Ocasional structural breaks and long memory', *UCSD Economics Discussion Paper* 98-03.

Kilian, L. and T. Zha (2002), 'Quantifying the uncertainty about the half-life of deviations from PPP', *Journal of Applied Econometrics,* 17, 107-125.

Kilian, L. and M.P. Taylor (2003), 'Why is it so difficult to beat the random walk for exchange rates?', *Journal of International Economics,* 60, 85-107

Koop, G., M.H. Pesaran and S.M. Potter (1996), 'Impulse response analysis in nonlinear multivariate models', *Journal of Econometrics,* 74, 119-147

Lothian, J.R. and M.P. Taylor (1996), 'Real exchange rate behavior: the recent float from the perspective of the past two centuries'. *Journal of Political Economy,* 104, 488-509.

Lothian, J.R., and M.P. Taylor (2004), 'Real exchange rates over the past two centuries: How important is the Harrod-Balassa-Samuelson effect?'.

Michael, P., A.R. Nobay and D.A. Peel (1997), 'Transactions costs and nonlinear adjustment in real exchange rates: an empirical investigation', *Journal of Political Economy*, 105, 862-879.

Murray, C. and D. Papell (2002), 'The purchasing power parity persistence paradigm', *Journal of International Economics,* 56, 1-19.

Ozaki, T. (1985), 'Nonlinear time series models and dynamical systems', in E.J. Hannan, P.R. Krishnaiah, and M.M. Rao eds., *Handbook of Statistics, V*, Amsterdam, Elsevier, 25-83.

Paya, I. and D.A. Peel (2003), 'On the relationship between nominal exchange rates and domestic and foreign prices', *Mimeo, Cardiff Business School.*

Paya, I. and D.A. Peel (2004a), 'An analysis of the small sample properties on nonlinear estimates', *Mimeo Lancaster University.*

Paya, I. and D.A. Peel (2004b), 'Nonlinear PPP under the gold standard', *Southern Economic Journal,* 71, 302-313.

Paya, I. and D.A. Peel (2004c), 'A new analysis of the determinants of the real dollar-sterling exchange rate: 1871-1994'. *Mimeo, Cardiff Business School.*

Paya, I. and D.A. Peel (2004d), ‘Temporal aggregation of an ESTAR process: Some implications for purchasing power parity adjustment’, *Mimeo Lancaster University.*

Paya, I., I.A. Venetis and D.A. Peel (2003), ‘Further evidence on PPP adjusment speeds: The case of effective real exchange rates and the EMS’, *Oxford Bulletin of Economics and Statistics*, 65, 421-438.

Taylor, A.M. (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, 473-498.

Taylor, M.P., D.A. Peel, and L. Sarno (2001), ‘Nonlinear mean-reversion in real exchange rates: Toward a solution to the purchasing power parity puzzles’, *International Economic Review*, 42, 1015-1042.

Terasvirta, T. (1994), ‘Specification, estimation and evaluation of smooth transition autoregressive models’, *Journal of the American Statistical Association,* 89, 208–218

Van Dijk, D., P.H. Franses, and P.H. Boswijk (2000), ‘Asymmetric and Common Absorption of Shocks in Nonlinear Autoregressive Models’. *Econometric Institute Research Report* EI2000-01/A

Venetis, I.A., I. Paya, I. and D.A. Peel (2002), ‘Do real exchange rates ’mean revert ’ to productivity? A nonlinear approach’, *Mimeo, Cardiff Business School.*

| Table 1. Power of unit root test against STAR | | | | |
|---|---|---|---|---|
| | | | KSS | Kiliç |
| $\gamma = 1$ | $se = 0.035$ | *sample* 10,000 | 1 | 1 |
| | | *sample* 350 | 0.355 | 0.652 |
| | $se = 0.01$ | *sample* 10,000 | 1 | 1 |
| | | *sample* 350 | 0.091 | 0.331 |
| $\gamma = 0.5$ | $se = 0.035$ | *sample* 10,000 | 1 | 1 |
| | | *sample* 350 | 0.204 | 0.560 |
| | $se = 0.01$ | *sample* 10,000 | 1 | 1 |
| | | *sample* 350 | 0.077 | 0.305 |

| Table 2. Nonlinear test in temporal aggregated ESTAR | | | | | |
|---|---|---|---|---|---|
| | | | KSS | Kiliç | Both |
| Aggregation i=12 (annual aggregation) | | | | | |
| $\gamma = 1$ | $se = 0.035$ | *sample* 10,000 | 1 | 1 | 1 |
| | | *sample* 200 | 0.991 | 1 | 0.991 |
| | $se = 0.01$ | *sample* 10,000 | 1 | 1 | 1 |
| | | *sample* 200 | 0.680 | 0.601 | 0.441 |
| $\gamma = 0.5$ | $se = 0.035$ | *sample* 10,000 | 1 | 1 | 1 |
| | | *sample* 200 | 0.972 | 0.964 | 0.961 |
| | $se = 0.01$ | *sample* 10,000 | 1 | 1 | 1 |
| | | *sample* 200 | 0.401 | 0.395 | 0.198 |
| Aggregation i=3 (quarterly aggregation) | | | | | |
| $\gamma = 1$ | $se = 0.035$ | *sample* 40,000 | 1 | 1 | 1 |
| | | *sample* 120 | 0.308 | 0.254 | 0.169 |
| | $se = 0.01$ | *sample* 40,000 | 1 | 1 | 1 |
| | | *sample* 120 | 0.108 | 0.135 | 0.054 |
| $\gamma = 0.5$ | $se = 0.035$ | *sample* 40,000 | 1 | 1 | 1 |
| | | *sample* 120 | 0.195 | 0.189 | 0.091 |
| | $se = 0.01$ | *sample* 40,000 | 1 | 1 | 1 |
| | | *sample* 120 | 0.090 | 0.122 | 0.046 |
| Aggregation i=3 (monthly aggregation) | | | | | |
| $\gamma = 0.3$ | $se = 0.024$ | *sample* 10,000 | 1 | 1 | 1 |
| | | *sample* 360 | 0.423 | 0.359 | 0.271 |
| Aggregation i=2 (montly aggregation) | | | | | |
| $\gamma = 0.4$ | $se = 0.028$ | *sample* 10,000 | 1 | 1 | 1 |
| | | *sample* 360 | 0.340 | 0.450 | 0.230 |

Table 3a. Results for simulated aggregated data of ESTAR model

| | | | | |
|---|---|---|---|---|
| True DGP: | | $y_t = e^{-\gamma y_{t-1}^2} y_{t-1} + u_t$ | | |
| Estimated model: | | $y_t^* = a + B(L) y_{t-1}^* e^{-\gamma (y_{t-1}^* - a)^2}$ | | |
| | Aggregation i=12 (annual aggregation) | | | |
| | $\gamma = 1$ | | $\gamma = 1$ | |
| | $se = 0.035$ | | $se = 0.01$ | |
| | $sample 10,000$ | $sample$ 200 | $sample 10,000$ | $sample$ 200 |
| $Mean\ \widehat{\gamma}$ | 4.50 | 5.00 | 7.62 | 10.50 |
| $sd\ \widehat{\gamma}$ | 0.45 | 3.85 | 0.70 | 6.80 |
| $t(\widehat{\gamma})$ | 1.000 | 0.240 | 1.000 | 0.370 |
| $R^2$ | 0.60 | 0.60 | 0.86 | 0.85 |
| $se$ | 0.077 | 0.077 | 0.025 | 0.025 |
| $LM\ Arch$ | 1.000 | 0.183 | 0.300 | 0.070 |
| $AR(2)$ | 1.000 | 0.750 | 1.000 | 0.860 |
| $AR(3)$ | 0.995 | 0.095 | 0.990 | 0.120 |
| $AR(4)$ | 0.220 | 0.075 | 0.270 | 0.065 |
| $AR(5)$ | 0.070 | 0.070 | 0.060 | 0.060 |
| | $\gamma = 0.50$ | | $\gamma = 0.50$ | |
| | $se = 0.035$ | | $se = 0.01$ | |
| | $sample 10,000$ | $sample$ 200 | $sample 10,000$ | $sample$ 200 |
| $Mean\ \widehat{\gamma}$ | 2.78 | 3.30 | 4.07 | 5.95 |
| $sd\ \widehat{\gamma}$ | 0.24 | 2.00 | 0.38 | 4.58 |
| $t(\widehat{\gamma})$ | 1.000 | 0.360 | 1.000 | 0.340 |
| $R^2$ | 0.69 | 0.69 | 0.90 | 0.89 |
| $se$ | 0.082 | 0.082 | 0.026 | 0.026 |
| $LM\ Arch$ | 0.995 | 0.010 | 0.157 | 0.058 |
| $AR(2)$ | 1.000 | 0.770 | 1.000 | 0.840 |
| $AR(3)$ | 0.996 | 0.120 | 1.000 | 0.123 |
| $AR(4)$ | 0.220 | 0.080 | 0.290 | 0.067 |
| $AR(5)$ | 0.070 | 0.070 | 0.077 | 0.043 |

Notes: $sd$ denotes standard deviation of coefficient $\gamma$. $t(\widehat{\gamma})$ denotes ratio of significant $\gamma$ parameter where empirical significance level is obtained through Monte Carlo. $se$ denotes standard error of equation. $LM\ Arch$ is the ratio of rejection of the Lagrange Multiplier test for ARCH in the residuals. $AR(p)$ denotes ratio of significant autoregressive term of order $p$.

Table 3b. Results for simulated aggregated data of ESTAR model

| | | | | |
|---|---|---|---|---|
| True DGP: | $y_t = e^{-\gamma y_{t-1}^2} y_{t-1} + u_t$ | | | |
| Estimated model: | $y_t^* = a + B(L) y_{t-1}^* e^{-\gamma (y_{t-1}^* - a)^2}$ | | | |
| | Aggregation i=3 (quarterly aggregation) | | | |
| | $\gamma = 1$ | | $\gamma = 1$ | |
| | $se = 0.035$ | | $se = 0.01$ | |
| | *sample* 40,000 | *sample* 120 | *sample* 40,000 | *sample* 120 |
| $Mean\ \widehat{\gamma}$ | 2.14 | 3.50 | 2.37 | 13.80 |
| $sd\ \widehat{\gamma}$ | 0.09 | 3.00 | 0.17 | 19.50 |
| $t(\widehat{\gamma})$ | 1.000 | 0.390 | 1.000 | 0.270 |
| $R^2$ | 0.86 | 0.85 | 0.96 | 0.92 |
| $se$ | 0.047 | 0.047 | 0.013 | 0.013 |
| $LM\ Arch$ | 0.584 | 0.092 | 0.128 | 0.085 |
| $AR(2)$ | 1.000 | 0.460 | 1.000 | 0.510 |
| $AR(3)$ | 1.000 | 0.090 | 1.000 | 0.110 |
| $AR(4)$ | 0.370 | 0.066 | 0.350 | 0.083 |
| $AR(5)$ | 0.055 | 0.055 | 0.077 | 0.066 |
| | $\gamma = 0.50$ | | $\gamma = 0.50$ | |
| | $se = 0.035$ | | $se = 0.01$ | |
| | *sample* 40,000 | *sample* 120 | *sample* 40,000 | *sample* 120 |
| $Mean\ \widehat{\gamma}$ | 1.10 | 2.20 | 1.20 | 11.70 |
| $sd\ \widehat{\gamma}$ | 0.05 | 2.54 | 0.09 | 19.25 |
| $t(\widehat{\gamma})$ | 1.000 | 0.375 | 1.000 | 0.240 |
| $R^2$ | 0.90 | 0.88 | 0.97 | 0.92 |
| $se$ | 0.048 | 0.047 | 0.014 | 0.014 |
| $LM\ Arch$ | 0.310 | 0.070 | 0.127 | 0.087 |
| $AR(2)$ | 1.000 | 0.510 | 1.000 | 0.490 |
| $AR(3)$ | 1.000 | 0.100 | 1.000 | 0.090 |
| $AR(4)$ | 0.350 | 0.050 | 0.380 | 0.062 |
| $AR(5)$ | 0.070 | 0.070 | 0.068 | 0.078 |

Notes: see notes in table 1

| Table 3c. Results for simulated aggregated data of ESTAR model | | |
|---|---|---|
| True DGP: | | $y_t = e^{-\gamma y_{t-1}^2} y_{t-1} + u_t$ |
| Estimated model: | | $y_t^* = a + B(L) y_{t-1}^* e^{-\gamma (y_{t-1}^* - a)^2}$ |
| | Aggregation i=3 (monthly aggregation) | |
| | $\gamma = 0.3$ | |
| | $se = 0.024$ | |
| | $sample$ $10,000$ | $sample$ 360 |
| $Mean\ \widehat{\gamma}$ | 0.71 | 1.09 |
| $sd\ \widehat{\gamma}$ | 0.08 | 0.75 |
| $t(\widehat{\gamma})$ | 1.000 | 0.390 |
| $R^2$ | 0.95 | 0.93 |
| $se$ | 0.033 | 0.033 |
| $LM\ Arch$ | 0.114 | 0.100 |
| $AR(2)$ | 1.000 | 0.890 |
| $AR(3)$ | 0.950 | 0.110 |
| $AR(4)$ | 0.130 | 0.080 |
| $AR(5)$ | 0.052 | 0.055 |

Notes: see notes in table 1

| Table 3d. Results for simulated aggregated data of ESTAR model | | |
|---|---|---|
| True DGP: | | $y_t = e^{-\gamma y_{t-1}^2} y_{t-1} + u_t$ |
| Estimated model: | | $y_t^* = a + B(L) y_{t-1}^* e^{-\gamma (y_{t-1}^* - a)^2}$ |
| | Aggregation i=2 (monthly aggregation) | |
| | $\gamma = 0.4$ | |
| | $se = 0.028$ | |
| | $sample$ $10,000$ | $sample$ 360 |
| $Mean\ \widehat{\gamma}$ | 0.67 | 1.01 |
| $sd\ \widehat{\gamma}$ | 0.08 | 0.76 |
| $t(\widehat{\gamma})$ | 1.000 | 0.390 |
| $R^2$ | 0.95 | 0.93 |
| $se$ | 0.033 | 0.033 |
| $LM\ Arch$ | 0.126 | 0.084 |
| $AR(2)$ | 1.000 | 0.700 |
| $AR(3)$ | 0.590 | 0.080 |
| $AR(4)$ | 0.070 | 0.050 |
| $AR(5)$ | 0.050 | 0.045 |

Notes: see notes in table 1

Table 4. Results from ESTAR models of bilateral dollar real exchange rates.
Montly observations 1973-2001.

| | $\hat{\delta}_0$ | $\hat{\beta}_1$ | $\hat{\beta}_2$ | $\hat{\gamma}$ | $s$ | $p\ AR(2)$ | Half-life to 10% shock |
|---|---|---|---|---|---|---|---|
| FRF | -0.025 | 1.037 | $\beta_2 = 0$ | 0.779 | 0.031 | 0.55 | 12 |
| | (0.031) | (0.022) | | (0.313) | | | |
| BEF | 0.005 | 1.018 | $\beta_2 = 0$ | 0.331 | 0.033 | 0.46 | 25 |
| | (0.048) | (0.020) | | (0.185) | | | |
| DEM | -0.027 | 1.033 | $\beta_2 = 0$ | 0.625 | 0.033 | 0.27 | 13 |
| | (0.036) | (0.021) | | (0.248) | | | |
| ITL | -0.045 | 1.017 | $\beta_2 = 1 - \beta_1$ | 0.336 | 0.030 | 0.15 | 36 |
| | (0.043) | (0.022) | | (0.194) | | | |
| JPY | 0.479 | 1.105 | $\beta_2 = 1 - \beta_1$ | 0.155 | 0.033 | 0.05 | 40 |
| | (0.059) | (0.053) | | (0.082) | | | |
| NLG | 0.041 | 1.022 | $\beta_2 = 0$ | 0.481 | 0.033 | 0.48 | 18 |
| | (0.046) | (0.022) | | (0.236) | | | |
| GBP | 0.109 | 1.094 | $\beta_2 = 0$ | 0.595 | 0.031 | 0.16 | 24 |
| | (0.059) | (0.069) | | (0.361) | | | |

Notes: Numbers in parentheses are Newey-West standard error estimates. $s$ denotes the residuals standard error. $pAR(2)$ denotes p-value of second autoregressive term in the ESTAR estimation.
Source: Table from Venetis et al. (2002)

Table 5. Results from ESTAR models of bilateral dollar real exchange rates.
Quarterly observations 1973-1998

| | $\hat{\delta}_0$ | $\hat{\beta}_1$ | $\hat{\beta}_2$ | $\hat{\gamma}$ | $s$ | Half-life to 10% shock |
|---|---|---|---|---|---|---|
| FRF | 0.095 | 1.32 | $\beta_2 = 1 - \beta_1$ | 0.964 | 0.047 | 36 |
| | (0.033) | (0.096) | | (0.152) | | |
| DEM | 0.096 | 1.233 | $\beta_2 = 1 - \beta_1$ | 0.794 | 0.053 | 36 |
| | (0.036) | (0.099) | | (0.125) | | |
| CAN | 0.00 | 1.181 | $\beta_2 = 1 - \beta_1$ | 0.706 | 0.019 | 40 |
| | | (0.078) | | (0.043) | | |
| ITL | 0.00 | 1.154 | $\beta_2 = 1 - \beta_1$ | 0.909 | 0.054 | 36 |
| | | (0.113) | | (0.247) | | |
| JPY | 0.00 | 1.350 | $\beta_2 = 1 - \beta_1$ | 0.725 | 0.057 | 40 |
| | | (0.103) | | (0.094) | | |
| SW | 0.00 | 1.292 | $\beta_2 = 1 - \beta_1$ | 0.724 | 0.059 | 40 |
| | | (0.099) | | (0.139) | | |
| GBP | 0.00 | 1.144 | $\beta_2 = 1 - \beta_1$ | 1.069 | 0.052 | 36 |
| | | (0.103) | | (0.324) | | |

Notes: Numbers in parentheses are Newey-West standard error estimates. $s$ denotes the residuals standard error. Source: Table from Kilian and Taylor (2003)

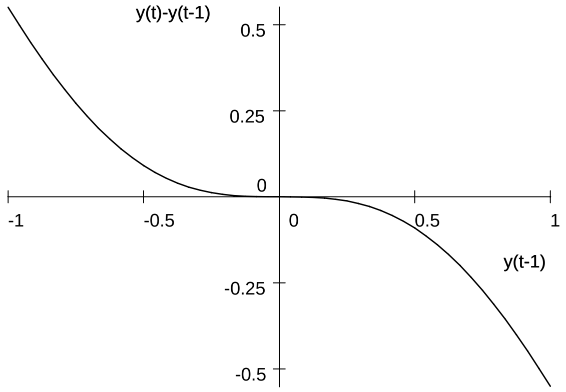
Table 6. Results from ESTAR models of bilateral dollar real exchange rates. Annual observations

| | $\hat{\delta}_0$ | $\hat{\beta}_1$ | $\hat{\beta}_2$ | $\hat{\gamma}$ | $s$ | Half-life to 10% shock |
|---|---|---|---|---|---|---|
| Dollar/FrF | -0.083 | 1.12 | $\beta_2 = 1 - \beta_1$ | 4.03 | 0.076 | 36 |
| 1804-1992 | (0.025) | (0.15) | | (1.54) | | |
| Dollar/Pound | -0.210 | 1.18 | $\beta_2 = 1 - \beta_1$ | 2.43 | 0.069 | 48 |
| 1792-1992 | (0.019) | (0.069) | | (0.54) | | |
| Dollar/DM | -0.033 | 1.09 | $\beta_2 = 1 - \beta_1$ | 2.52 | 0.095 | 48 |
| 1795-1913 | (0.032) | (0.08) | | (0.60) | | |

Notes: Numbers in parentheses are Newey-West standard error estimates. $s$ denotes the residuals standard error

Table 7: Estimated half-lives of shocks in months from Temporally aggregated Data: True DGP $y_t = e^{-\gamma y_{t-1}^2} + u_t$ where $u_t = N(0, se)$

| Line | Shock: | 5% Months | | | 20% Months | | | 30% Months | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Temporal Aggregation Large Sample | | 25% | 50% | 80% | 25% | 50% | 80% | 25% | 50% | 80% |
| 1 | True DGP $\gamma = 1, se = 0.035$ | 7 | 18 | 45 | 6 | 16 | 42 | 4 | 13 | 38 |
| 2 | i=3: $\gamma = 2.14, se = 0.047$ | 12 | 27 | 66 | 7 | 21 | 60 | 4 | 14 | 51 |
| 3 | i=12: $\gamma = 4.5, se = 0.077$ | 20 | 42 | 97 | 12 | 27 | 84 | 6 | 14 | 72 |
| 4 | True DGP $\gamma = 0.4, se = 0.028$ | 8 | 19 | 47 | 6 | 16 | 42 | 3 | 12 | 39 |
| 5 | i=2 $\gamma = 0.67, se = 0.035$ | 8 | 22 | 56 | 6 | 18 | 51 | 4 | 14 | 46 |
| 6 | True DGP $\gamma = 0.3, se = 0.024$ | 7 | 16 | 37 | 4 | 13 | 31 | 3 | 10 | 29 |
| 7 | i=3 $\gamma = 0.71, se = 0.035$ | 9 | 23 | 57 | 6 | 18 | 52 | 4 | 15 | 47 |
| Temporal Aggregation Small Sample | | | | | | | | | | |
| 8 | True DGP $\gamma = 1.65, se = 0.035$ | 6 | 14 | 34 | 4 | 11 | 31 | 2 | 8 | 27 |
| 9 | i=3: $\gamma = 3.5, se = 0.047$ | 11 | 24 | 54 | 6 | 16 | 48 | 3 | 12 | 42 |
| 10 | i=12: $\gamma = 5, se = 0.077$ | 21 | 40 | 96 | 12 | 26 | 82 | 3 | 14 | 72 |
| 11 | True DGP $\gamma = 0.7, se = 0.028$ | 4 | 11 | 30 | 3 | 8 | 27 | 2 | 7 | 23 |
| 12 | i=2: $\gamma = 1, se = 0.035$ | 7 | 19 | 46 | 4 | 15 | 39 | 3 | 12 | 35 |
| 13 | True DGP $\gamma = 0.5, se = 0.024$ | 4 | 9 | 22 | 3 | 5 | 20 | 2 | 4 | 16 |
| 14 | i=3: $\gamma = 1, se = 0.035$ | 6 | 16 | 44 | 5 | 14 | 41 | 3 | 10 | 35 |




Deterministic plot of $\Delta y$ (vertical axis), $y_{t-1}$(horizontal axis) from ESTAR with $\gamma = 0.8$.