

Nonlinear PPP Deviations: A Monte Carlo Investigation of their Unconditional Half-Life

Ming Chien Lo[†]
Department of Economics
St. Cloud State University

November 9, 2006

Abstract:

Recent research has generated support to the notion that the real exchange rate adjustment is nonlinear and that the PPP half-life is faster than the puzzling 3 to 5 years based on linear models. While different nonlinear models survive the specification tests against linear ones, there is little consensus on which specific threshold-type model outperforms the others in the family. In this paper, a Monte Carlo study is designed to address the issue and the findings support that the MR-LSTAR process is the most likely suspect that generates the puzzle.

JEL Code: C15, C22, F31

Keywords: Generalized Impulse Response Functions, Monte Carlo, Nonlinear Autoregressive Model, Purchasing Power Parity, Real Exchange Rates

[†] Department of Economics SH386, St. Cloud State University, 720 Fourth Avenue South, St. Cloud, MN 56301-4498. Telephone: (320) 308-0143. Fax: (320) 308-2228. Email: mclo@stcloudstate.edu.

Acknowledgement

An earlier version of the paper titled “Bias in Measuring the Persistence of Nonlinear PPP Deviations” was presented at the 13th Annual Symposium of the Society of Nonlinear Dynamics and Econometrics held at the City University, London, United Kingdom and the 15th Fall meeting of the Midwest Econometric Group, and the economic seminar series at the University of Manitoba. The author thanks Bruce Mizrach, Phil Rothman and Dick van Dijk and the organizers of both conferences, Squib Jafarey and Guilia Iori in particular, for the opportunity and the audiences in the session for their comments. The author also thanks Robert-Paul Berben, Ryan Compton, Jennifer Goulet, James Morley, David Peel and Dick van Dijk. Any error is the sole responsibility of the author.

1. Introduction

The theory of purchasing power parity (PPP) has long been providing a simple and convenient testing field to examine time series models. After years of cycles of attempts, frustration and marginal success, a general consensus emerges: (i) more powerful unit root tests uncover some evidence that real exchange rates are stationary, (ii) the estimated half-life of a PPP deviation ranges from 3 to 5 years (see Rogoff 1996) and (iii) these estimates are nevertheless too large. Recent research focuses on two directions that attempt to examine whether they are results of aggregation bias (e.g. Chen and Engel, 2005 and Imbs et al, 2005) and/or model mis-specification bias (e.g. Bec, Ben Salem and Carrasco, 2004, Imbs, 2003, Lo and Zivot, 2001, Michael, Nobay and Peel, 1997, Obstfeld and Taylor, 1997, O'Connell and Wei, 2002, Sarantis, 1999.) Literature motivated by the latter argument is generally founded on a theoretical model assuming transaction costs of commodity arbitrage; the resulting data generating process, the authors argue, is nonlinear. For a more detailed and up-to-date overview of the development, see Taylor and Sarno (2003) and Sarno (2003).

As a concept taught in the classroom, PPP is usually introduced in a world without transaction costs. Since a costless world is unrealistic outside the classroom, some researchers argue that a pair of transaction-cost bands around the equilibrium of real exchange rates should be incorporated in the econometric specification. The nonlinear time series models used often belong to the family of threshold-type.¹ By imposing such models, we may find a much faster speed of equilibrium adjustment outside the bands when market forces are effective.

¹ In contrast to, say, Markov-switching time series models.

Many nonlinear studies directly compare estimates based on the data in the “outside-band” regimes, i.e. sub-set of the full sample, with the estimate based on the full sample, like those reported in Rogoff (1996). As noted in the important contribution by Koop, Pesaran and Potter (1996) and Potter (2000), hereafter KPP, the dynamics of the impulse responses as well as the way to evaluate them can change drastically if we move from a linear model to a nonlinear one. Given a linear autoregressive model, the size and the initial conditions generally do not alter the characteristics of the impulse responses. The half-life estimate based on a linear impulse response function is thus constant regardless of the assumption set up for the perturbation of the function. Many nonlinear autoregressive models by definition assume changes in the autoregressive dynamics that depends on the properties of shocks and initial conditions. As a result, the uniform characteristics of the impulse responses observed under a linear model no longer holds in this case. In this regard, the meaning of half-life for a nonlinear model becomes complicated. Ignoring this difference between a linear and a nonlinear impulse response function may result in confusing inferences.

Kiliç (2005) and Shantini (2006), each works independently, attempt to develop a half-life measure appropriate for nonlinear models to overcome the shortcomings of conventional computations. Their contributions are important but the scope of their work is limited to specific nonlinear models. Various model specification tests against linearity and/or unit root have also been motivated in Teräsvirta and Anderson (1992), Bec, Ben Salem and Carrasco (2004, 2005) and Hansen (1997). With the exception of Bec, Ben

Salem and Carrasco, these studies barely touch the issue of impulse response functions in the context of nonlinear models.²

This paper attempts to fill the void in two areas. First, unlike previous studies focusing on only one or a few models, it examines all major nonlinear models that have been used for the challenge of the PPP persistence puzzle. In particular, I want to narrow down the specific nonlinear model(s) that is most likely responsible in generating the puzzling, *unconditional* half-life of 3-5 years as reported in Rogoff (1996). Why should we indirectly examine models in this way and not use available procedures for specification testing? Each of the previous studies found support for their specific type of threshold autoregressive models which differ mainly in the regime-switching functions; to name a few, discrete switching in Obstfeld and Taylor's (1997), exponential smooth transition in Michael, Nobay and Peel (1997) and mirrored logistic smooth transitions in Bec, Ben Salem and Carrasco (2004). Since these models are not all nested, it is hard to test which gives a better fit to the data. Moreover, specification tests involve pair-wise contests between two models; work becomes tedious when there are three or four nonlinear models under examination.

Therefore, the second goal of this paper is to integrate various nonlinear time series models with KPP's generalized impulse response function (GIRF) to study nonlinear PPP deviations and equilibrium adjustment. Monte Carlo experiments are used to draw inferences regarding the unconditional bias in linear half-life estimate. Unlike Berben

² Bec, Ben Salem and Carrasco (2004, 2005) applied KPP's generalized impulse response function on their specific threshold autoregressive model for multiple regimes on real exchange rates. They also make use of the *absorption* measure developed by van Dijk et al (2005). This measure, however, is not related to the half-life issue here.

(2000) and some others, this study does not simply apply the GIRF. Section 3 will discuss the specific contribution in developing a relative persistence measure that is more appropriate in the current context. Section 4 will apply this methodology on the U.S.-G6 real exchange rates.

2. Regime-Specific versus Unconditional Impulse Responses

Rogoff (1996) generalized that the half-life estimates for PPP in the literature ranged from 3 to 5 years. The studies surveyed by him were mostly based on a linear modeling framework. Authors who adopt a nonlinear framework generally regard this as a bias due to model mis-specification. Regardless of the specific nonlinear model used, many studies, most notably Obstfeld and Taylor (1997), compare the persistence estimate from a linear model, e.g. an AR(1) and the persistence estimate for a specific regime from a nonlinear model.³ Their analyses are best illustrated in Figure 1. Suppose a three-regime threshold model represents the true data generating process (hereafter, DGP) of the logarithm of the real exchange rate. Any data point within the bands represents a small equilibrium PPP deviation. Such a small deviation means that the potential gain from commodity arbitrage is smaller than the transaction costs approximated by the bands and thus no equilibrium process will take place to adjust the real exchange rate. When the real exchange rate lays in the outside-band regime, the deviation is considered large. In this case, arbitrage becomes profitable and we should expect the exchange rate to converge back towards the inside-band regime. The detailed specifications vary from model to model. The example used in

³ In this paper, I interchangeably use the term AR(1) and the equivalent error-correction form of the same order.

Figure 1 is consistent with Obstfeld and Taylor's (1997) threshold autoregressive (TAR) model using an AR(1) model to capture outside-band convergence and a random walk model to capture the inside-band absence of equilibrium adjustment. To be specific, denote q_t the logarithm of the real exchange rate, the general model is of the form:

$$\Delta q_t = \begin{cases} \phi_{01} + \phi_{11}q_{t-1} + \varepsilon_{1t}, & \text{if } q_{t-1} > c, \\ \varepsilon_{2t}, & \text{if } |q_{t-1}| \leq c, \\ \phi_{03} + \phi_{13}q_{t-1} + \varepsilon_{3t}, & \text{if } q_{t-1} < -c. \end{cases} \quad (1)$$

If $\phi_{01} = \phi_{03} = 0$ and $\phi_{11} = \phi_{13} = \phi_{TAR}$, the model becomes an Equilibrium-TAR (hereafter, EQTAR) model which allows for convergence towards the equilibrium level (or zero in our example.) If $\phi_{01} = c(1 - \phi_{TAR})$, $\phi_{03} = c(-1 + \phi_{TAR})$ and $\phi_{11} = \phi_{13} = \phi_{TAR}$, the model becomes a Band-TAR model which allows for convergence towards the bands. In other words, the EQTAR model and the Band-TAR model each is nested in the general model (1), but neither of them is nested in the other. If $\phi_{11} \neq \phi_{13}$, the speed of adjustment above and below the inside-band regime are different; we often describe such a three-regime model *asymmetric*. Obstfeld and Taylor (1997) use the regime-specific estimate for ϕ_{TAR} and compute the half-life measure $\rho_{TAR} = \ln 0.5 / \ln(1 + \phi_{TAR})$. This measure is then compared with the half-life estimate based on a linear AR(1) model in error correction form. Taylor (2001) later followed this direction and mathematically determined the half-life bias.

As long as the estimate for ϕ_{TAR} is negative and significantly smaller than the estimate from linear models, it is comforting to know that market PPP adjustment is faster than once perceived. Yet, the estimate for ϕ_{TAR} is conditional on the data of a sub-sample and regime-specific. From an empirical standpoint, we do not exactly know the actual improvement (or

reduction) in the persistence of a PPP deviation in terms of the unconditional estimate of half-life based on linear models, which are the focus of Rogoff (1996).⁴ If these estimates are biased, they are the results of the true DGP in the full sample; any direct comparison between these estimates and the half-life estimates computed from ϕ_{TAR} does not necessarily explain the whole story regarding the model-specification bias.

Assume that the argument against linear modeling is correct and that the true DGP follows the specification of, say, a Band-TAR model. If we mistakenly apply an AR(1) model on the data, the estimate for the autoregressive coefficient, denoted ϕ_{AR} , will lie between ϕ_{TAR} and zero. This also implies that $\rho_{TAR} < \rho_{AR}$ where $\rho_{AR} = \ln 0.5 / \ln(1 + \phi_{AR})$. A problem of this comparison arises. ρ_{TAR} is an appropriate persistence estimate only if we restrict our analysis to a very specific condition. In particular, this is a valid measure only when we consider a positive shock when the real exchange rate initially lies above the upper band *or* a negative shock when the real exchange rate initially lies below the lower band. When the signs of the shock or the initial conditions are different, the cross-regime dynamics may detour our adoption of a simple half-life measure. In the outside-band regime, if a shock away from the bands is not strong enough to bring the dynamics back across the threshold, the true and relevant persistence measure is ρ_{TAR} . In this case, ρ_{AR} is upward-biased. In the inside-band regime, if an initial shock is not big enough to bring the series across the threshold, then the true and relevant persistence measure is no longer ρ_{TAR} —since the dynamics follows random walk in this regime, all shocks have permanent

⁴ Half-life estimated for linear models are based on the information of the full sample. It can be described as unconditional (on regimes or sub-sample.)

effect. As a result, ρ_{AR} is downward-biased. What complicates the matter even more is a shock that is large enough to bring in cross-regime dynamics. In this case, the direction of the bias of ρ_{AR} is an empirical issue. In short, from an empirical standpoint we may not want to simply compare a regime-specific estimate ρ_{TAR} with a full in-sample or unconditional estimate for ρ_{AR} . If we look for the general or non-regime-specific in-sample degree of persistence in the data given that the true DGP is nonlinear, we must make use of the joint distribution of the shocks and the history.

KPP have pointed out the properties of shock and history dependence in nonlinear impulse response analysis. They developed a procedure to analyze the degree of persistence for various given assumptions on shocks and history. To illustrate the criticism based on their analysis, I conducted a Monte Carlo experiment given equation (1) with both EQTAR and Band-TAR specifications. The DGP in this experiment assumes $\phi_{11} = \phi_{13} = -0.3$ and $c = 1.5$. I allowed the standard deviation to vary in the inside-band and the outside-band regimes; the variation ranged from 0.5 to 3.0 with a 0.5 interval. The sample size for each simulation is 500, and 10000 trials were simulated for each set of parametric values. Given the true DGP is nonlinear, a linear AR(1) model was used in each trial for fitting the data. The Monte Carlo mean as well as the standard errors from the AR(1) model are reported in the two panels in Table 1.

The direction of bias on the autoregressive coefficient is uniform for the Band-TAR model, although the size of the bias varies quite a little from positive 0.6 to 2.2. The results for the EQTAR model are more complicated. While the majority of the mean coefficient estimates shows an upward bias, the bias is not very significant as the standard deviation

increases to a level comparable to the size of the bands. In a few cases, there is even slight downward bias when the standard deviation becomes very large. This exercise illustrates that in a small finite sample the empirical bias can be very different from case to case on the size of the shocks. A direct implication of this exercise is that we cannot consider identify the difference between the ϕ_{AR} and the ϕ_{TAR} estimates or between the ρ_{AR} and the ρ_{TAR} as bias without considering the size of the shocks.

A second criticism based on KPP is that impulse responses under a nonlinear framework necessarily involve cross-regime dynamics in the following sense. By definition, a generalized impulse response function is the difference between a simulated series given a specific initial shock and another simulated series created under the same assumption without the shock. The size of the shock and the initial condition in general do not matter for any persistence analysis based on a linear model. The opposite is true for nonlinear models. Figure 2 illustrates some variations when the functions are based on an EQTAR model. By varying the initial conditions, the shock of the same size of 2 units can lead to different half-live measure. Both Figure 2(a) and 2(b) indicate a half-life of approximately 3.8 periods. However, Figure 2(c) indicates that the half-life can be infinity.

Worse, under a linear framework, the assumption of additional future shocks that come after the initial shock is irrelevant because there is no cross-regime dynamics. KPP argue eloquently that this is not the same for nonlinear models. If the initial shock is specific enough to trigger cross-regime dynamics, the impulse response function can become harder to predict when future shocks are not set to zero. This is because the two simulated series will exhibit very mixed and very complex dynamic behavior in this case. Figure 2(d) shows

a single simulation with randomized future shocks (after period 0). Given that the initial shock is positive and of the size of 2, the simulated responses cross the half-life line of 1 least four times, and overall, there is no reversion to the initial condition after 30 periods. It is worth noticing that they also cross the lower half-life boundary of -1 multiple times. In short, taking randomized future shocks into consideration renders us to doubt the appropriateness of using conventional half-life as a persistence measure. Berben (2000) investigated the nonlinear impulse response for real exchange rate data. However, his analysis still used half-life to examine the degree of persistence.

Besides Band-TAR and EQTAR models, I also examined two of the smooth transition autoregressive (STAR) models, which are equally if not more popular in applied literature. Bec, Ben Salem and Carrasco (2004) developed a multiple-regime logistic STAR (or MR-LSTAR) model:⁵

$$\Delta q_t = (\mu_1 + \varphi_1 q_{t-1})G_1(q_{t-1}, \gamma, \lambda) + (\mu_2 + \varphi_2 q_{t-1})G_2(q_{t-1}, \gamma, \lambda) + (\mu_3 + \varphi_3 q_{t-1})G_3(q_{t-1}, \gamma, \lambda) + \varepsilon_t \quad (2)$$

where

$$G_1(q_{t-1}, \gamma, \lambda) = \left[1 + \exp\left(\frac{\gamma}{\sigma_{q_{t-1}}^T}(q_{t-1} + \lambda)\right) \right]^{-1},$$

$$G_3(q_{t-1}, \gamma, \lambda) = \left[1 + \exp\left(-\frac{\gamma}{\sigma_{q_{t-1}}^T}(q_{t-1} - \lambda)\right) \right]^{-1},$$

$$G_2(q_{t-1}, \gamma, \lambda) = 1 - G_1(q_{t-1}, \gamma, \lambda) - G_3(q_{t-1}, \gamma, \lambda).$$

The model is estimated by constrained maximum likelihood method, which is equivalent to first imposing values for γ and λ and then running an ordinary least square

⁵ The model in Bec, Ben Salem and Carrasco (2004) contains certain lags of the first difference of q_t on the right hand side.

regression. Note that $\sigma_{q_{t-1}}^T$ is a scaling factor, which is the sample standard deviation of the threshold variable q_{t-1} . In principle, this model allows for an asymmetric speed of adjustment. However, in practice, Bec, Ben Salem and Carrasco (2004) restrict that $\mu_1 = -\mu_3 = -\mu^*$ and $\varphi_1 = \varphi_3 = \varphi^*$, following the symmetry restrictions in Obstfeld and Taylor (1997). Their model outperforms the conventional exponential- and logistic- STAR models in terms of unit root rejection. The EQTAR model nests in this model if we set further restrictions: $\mu^* = \mu_2 = 0$, $\varphi_2 = 0$ and $\gamma \rightarrow \infty$.

The other STAR model is popularized by Michael, Nobay and Peel (1997) and has been extensively applied since their publication.⁶ The exponential STAR (or ESTAR) model is symmetric but does not nested with the TAR models:

$$\Delta q_t = (\mu_1 + \varphi_1 q_{t-1})F(q_{t-1}, \gamma, \lambda) + (\mu_2 + \varphi_2 q_{t-1})(1 - F(q_{t-1}, \gamma, \lambda)) + \varepsilon_t \quad (3)$$

where

$$F(q_{t-1}, \gamma, \lambda) = 1 - \exp\left[-\frac{\gamma}{\sigma_{q_{t-1}}^T} (q_{t-1} - \lambda)^2\right].$$

The ESTAR model is nested with neither the TARs and the MR-LSTAR. That said, Bec, Ben Salem and Carrasco (2004) shows that their mirrored logistic function can approximate an exponential function.

3. Mean Bias in Measuring Persistence under a Nonlinear Framework

KPP define a generalized impulse response function for a univariate model as the difference between two expected series of the variable, conditional on different

⁶ Most notable, O'Connell and Wei (2002) and Sarantis (1999).

assumptions on (i) history (or initial condition) $\Theta = (\theta_{t-i}, \theta_{t-i+1}, \dots, \theta_{t-1})$, (ii) shock(s) of interest $\Delta = (\delta_t, \delta_{t+1}, \dots, \delta_{t+j})$ and (iii) randomized shocks $\mathbf{V} = (v_t, v_{t+1}, \dots, v_{t+j})$. Given a model $q_t = g(\bar{\mathbf{q}}_i, \varepsilon_t)$ where $\bar{\mathbf{q}}_i = (q_{t-i}, q_{t-i+1}, \dots, q_{t-1})$. A generalized impulse response function can then be defined as

$$I_q(k, \mathbf{V}, \Delta, \Theta) = E(q_{t+k} \mid \bar{\mathbf{q}}_i = \Theta, \varepsilon_t = v_t + \delta_t, \varepsilon_{t+1} = v_{t+1} + \delta_{t+1}, \dots, \varepsilon_{t+k} = v_{t+k} + \delta_{t+k}) - E(q_{t+k} \mid \bar{\mathbf{q}}_i = \Theta, \varepsilon_t = v_t, \varepsilon_{t+1} = v_{t+1}, \dots, \varepsilon_{t+k} = v_{t+k}). \quad (4)$$

A simple simulation of (4) for a linear model, which is consistent with the half-life computation, assumes $\Theta = \mathbf{0}$, $\delta_t = 1$ and $\delta_{t+k} = 0 \forall k \geq 1$. Beaudry and Koop (1994) use a more carefully designed simulation with two shocks of interest to derive their nonlinear impulse responses for the U.S. GDP, which imply positive shocks have permanent effect while negative shocks have only transitory effect.

For a linear model, the assumptions on the value of v_t 's do not matter. The intuition behind this is that there is no cross-regime dynamics under the model specification. KPP recognize that the same does not hold for nonlinear model; see Figure 2(d) for an example. In addition, a nonlinear impulse response function also depends on the assumptions of history and the properties of the shocks—most notably, the sign and the size; see Figure 2(a)-(c) for a few examples. KPP thus advocate the use of simulation with specific design to randomly draw history and shocks that will answer specific questions. By allowing a randomization of history and shocks from a interested set of values, we can simulate a distribution of $I_q(k, \mathbf{V}, \Delta, \Theta)$ over a range of k and make inferences.

To date, there are only a handful of studies that adopt the nonlinear simulation technique to investigate PPP deviations even though a much large number of studies have

used nonlinear threshold autoregressive models to evaluate PPP adjustments in the presence of transaction costs. The effort of these studies is admirable. Nevertheless, they continue to use *conventional* half-life estimates as a measure of persistence. Kiliç (2005) and Taylor, Peel and Sarno (2001) are two notable exceptions that consider measures that are adjusted for nonlinearities. As illustrated in Figure 2, if we allow for various history and shocks, we may not only want to use the GIRF approach but also generate a measure of persistence different to the conventional half-life.

In this paper, I follow KPP and apply a similar simulation technique. In order to compare the results between the linear and nonlinear models, I assume the true DGP is nonlinear and simulate both a nonlinear impulse response function and a linear impulse response function based on the same Δ in order to make a reasonable comparison viable.⁷ Let $g(\bar{\mathbf{q}}_i, \varepsilon_t)$ follow an AR(1) model with a constant and one lag term and $g^*(\bar{\mathbf{q}}_i, \varepsilon_t)$ follow one of the nonlinear autoregressive models. The mean bias in the impulse responses due to model mis-specification is defined as

$$MBI(k, \mathbf{V}; \Delta_j, \Theta_j) \equiv I_q(k; \Delta_j) - \frac{\sum_{i=1}^n I_{q^*}(k, \mathbf{V}_i, \Delta_j, \Theta_j)}{n}. \quad (4)$$

In the Monte Carlo experiment, Δ_j and Θ_j are the j th set of shocks of interest and history drawn from specific distributions. The second term on the right hand side is the mean of the n simulated nonlinear impulse response functions with fixed Δ_j and Θ_j but randomized \mathbf{V}_i , where subscript i denotes the sub-trial in KPP's nonlinear impulse response simulation within each Monte Carlo simulation trial. The first term is the linear

⁷ History and randomized shocks are irrelevant to linear impulse response functions.

impulse response function, which is shock and history independent; hence, the notation is simplified. If $MBI(k, \mathbf{V}; \Delta_j, \Theta_j) > 0$ when the initial shock δ_t is *positive* or $MBI(k, \mathbf{V}; \Delta_j, \Theta_j) < 0$ when the shock is *negative*, we can infer that the linear model *on average* generates a more persistent impulse response at a particular point k for $k > 0$.

In each trial, I first randomly draw q_{t-1} as the initial condition from the empirical distribution. As heteroscedasticity is allowed in the TARs, a regime-specific initial shock δ_0 is drawn, while $\delta_{t+k} = 0 \forall k \geq 1$ is assumed. For the STAR model, the draw is from the full sample distribution of the estimated residuals. Randomized shocks v_t 's are also drawn in similar manner that may depend on the regime (again, for the TARs but not the STAR).

Two modifications of (4), however, are needed in the exercise.

Since the initial shock can be either positive or negative, inference from the sign of $MBI(k, \mathbf{V}; \Delta_j, \Theta_j)$ in (4) can be deceptive. To remedy this problem, I use a modified measure

$$MBI(k, \mathbf{V}; \Delta_j, \Theta_j) \equiv \left(I_q(k; \Delta_j) - \frac{\sum_{i=1}^n I_{q^*}(k, \mathbf{V}_i, \Delta_j, \Theta_j)}{n} \right) \times \text{sign}(\delta_0). \quad (5)$$

With the mean bias specified in (5), it is possible to derive an alternative measure of half-life bias. Suppose we ignorantly apply a linear autoregressive model to capture the dynamics of a data series, which is generated from a nonlinear process, and derive a linear impulse response function with the formula $(1 + \phi_{AR})^k$.

A reasonable half-life measure for the purpose of hypothesis testing can be found by solving for k numerically that satisfies the following condition:

$$(1 + \phi_{AR})^k - MBI_{k,\alpha} = 0.5. \quad (7)$$

where $MBI_{k,\alpha}$ denotes the value of the $MBI(k, \mathbf{V}; \Delta_j, \Theta_j)$ at the α -th percentile of the Monte Carlo simulation for each k .

The effort presented in this paper concerns the degree of persistence. Another aspect of nonlinear impulse response functions is absorption. In their important contribution, van Dijk, Franses and Boswijk (2005) define an absorption measure as the minimum horizon beyond which the difference between the distributions of the nonlinear impulse response functions becomes negligible. This definition holds regardless of the property of the distributions, which, suggested by KPP, indicate the persistence of the effect of a shock. This distinction between persistence and absorption is pointed out in van Dijk, Franses and Boswijk (2005).

4. Empirical Results

The monthly real exchange rate data are CPI-based from March 1973 to the most recent date available when this paper is being written.⁸ The U.S. dollar is used as the common currency against the currency of six other G-7 countries. The data are collected by the International Monetary Fund and available in the International Finance Statistics database.⁹ The data are in logarithm and demeaned.¹⁰ Note that I take the specification test results from previous studies for granted and proceed to the half-life analysis.

⁸ Canada and Japan: March 1973-November 2004; France and Italy: March 1973-December 1998; United Kingdom: March 1973-December 2004.

⁹ All price index data are CPIs and the nominal exchange rate is the monthly average rate for the amount of U.S. dollars for one unit of non-U.S. currency.

¹⁰ For the TARs and the MR-LSTAR, 10% of the highest and the lowest of the absolute value of the log real exchange rate are trimmed off for the selection of the value of the threshold.

4.1 Comparing Regime-Specific Estimates

Table 3 reports estimates from an AR(1) model¹¹ and the nonlinear models. The linear half-life estimates correspond very closely to what Rogoff (1996) reports, ranging mostly from 3.58 to 4.73 years with two outliers, Germany with 1.18 years and Canada 8.71 years.

When one compares the “outside-band” half-life for the EQTAR, the improvement is mixed. There is some reduction in the half-lives for the four exchange rates, ranging from 0.12 to 1.60 years. Contrary to the expectation, the estimates for the USD/CND and the USD/Yen exchange rates are larger. The average reduction is merely 0.42 years. The results for the Band-TAR model are more astonishing. Ignoring the two outliers, the range of half-life is 0.63 to 2.78 years; even for the two outliers sizable reduction is observed—0.70 years for the USD/CND exchange rate and 0.88 for the USD/DM exchange rate. A careful look at the table implies that 50% of a specific “outside-band” deviation will last less than a year for the USD/DM, the USD/Lira and the USD/£ exchange rates. Even for the USD/Franc and the USD/Yen exchange rates the half-lives are now shorter than 3 years.

Table 2 also reports the results for the ESTAR and the MR-LSTAR. When PPP is at its full force, the average reduction in half-life is only moderate for the ESTAR. The average half-life is 3.53, 0.73 years shorter than the linear estimate. This is slightly faster than the EQTAR estimates but slower than the Band-TAR estimates. Similar to the results for EQTAR, the USD/CND exchange rates has a longer half-life under this model.

The largest reduction in half-life is found under the MR-LSTAR. The reduction is apparent for all six data series and the size is the largest among all nonlinear models,

¹¹ All AR(1) models used in the paper contains a constant term.

ranging from 1.04 (USD/DM) to 3.94 (USD/Yen). The average half-life is a mere 1.63 years. However, we must be very cautious with the comparison here. φ^* from MR-LSTAR represents the faster speed of market adjustment; it is effective only when the deviation is very large as $G_1 = 1$ or $G_3 = 1$. But these conditions are satisfied fewer times than the threshold conditions of the TARs in the sample.¹² Although the comparison of the half-life estimates between the TARs and the MR-LSTAR is informative, it nevertheless oversimplifies the matter. This issue echoes the main theme of this paper that we need to compare the degree of persistence estimated from all models unconditionally.

4.2. Mean Bias of Impulse Responses

In the Monte Carlo experiment, m is set to 5,000 and n is set to 500 for each data series and each of the four nonlinear models (EQ-TAR, Band-TAR, ESTAR and MR-LSTAR.) First, I estimated the data using a nonlinear model. Then, the estimated coefficients and the distribution of the estimated residuals were used as the true parameters and true distribution for the simulation. In every trial, the simulated data were estimated by the true and the incorrect linear models, and linear impulse responses, nonlinear impulse responses and mean bias as defined in Section 3 were generated. A six-year horizon (72 months) was used for the impulse response simulation. I plotted the Monte Carlo median $MBI_{k,0.50}$ and two confidence intervals— $[MBI_{k,0.16}, MBI_{k,0.84}]$ and $[MBI_{k,0.05}, MBI_{k,0.95}]$ —for the four models in Figure 3.

¹² This is easy to observe. Interested readers are referred to the scattered plots of the empirical smooth transition functions in work by Michael, Nobay and Peel (1997) and O’Connell and Wei (2002). The density is almost always higher when the value of the function is less than one.

The results turn out to be slightly different from what are reported in the last section. As panel (a) shows, when the true DGP is of an EQTAR, the mean bias of applying an AR(1) is largely negative over all k horizons. With the exception of USD/DM, there is a tendency for the mean bias to decrease over time until a certain stable level. These estimates are also statistically significant for most of the time. The USD/£ exchange rate is the only series that has delayed (after 30 months) mean biases that are significantly non-zero. Panel (c) tells an opposite story. With the exception of the USD/Lira, the mean bias is mostly positive and significant. In the longer run, there is a tendency for the mean bias to decrease towards the zero line. Although the directions of bias are different, both panels (a) and (c) imply that it is highly unlikely that the EQTAR and the ESTAR are the true data generating process. If the opposite is true, then the linear model should be able to capture the *unconditional* half-life and the mean bias should be zero. As we observe significant non-zero bias over a horizon of 72 months for most of the cases, we can conclude that the EQTAR and the ESTAR should be ruled out from the suspect list. In addition, that the EQTAR model generates negative bias is not a result one can directly be inferred from those in Table 2. This reinforces the distinction between unconditional and conditional impulse responses.

Results shown in panel (b) and (d) support that both the Band-TAR and the MR-LSTAR are the true DGP. With the exception of the USD/£ (for the Band-TAR), the mean bias is zero over this 72-month horizon. Thus, the puzzling half-life estimated for a linear model is most likely generated by the two nonlinear models. It is not surprising that these two are the winners, as they are nested and the only difference between the two is the parametric restrictions in the Band-TAR. Moreover, since the MR-LSTAR has relatively

less restrictions, the findings for this model are more conclusive—even the mean bias for the USD/£ exchange rate is zero.

4.3 Median and Confidence Intervals for Half-Life Estimates under a Nonlinear Framework

Table 3 reports the median and the confidence intervals of the adjusted half-life estimates computed by solving for ρ_{NL} and the relevant critical value in equations (6) and (7). ρ_{NL} can be interpreted as the sample estimate of the true full-sample half-life under the hypothesis that the nonlinear process specifies the true model. The empirical confidence intervals $\rho_{NL,\alpha}$, as explained in Section 3, can then be used for hypothesis testing. The results in Table 3 agree with our analysis in the previous section. When the EQTAR is specified as the true model, the half-life estimates from the linear AR(1) model are smaller than the mean ρ_{NL} at mostly 1% significance level. In other words, if one believes in this model, the results suggest that Rogoff's half-life of 3-5 years is in fact an underestimation. If we ignore the outliers the USD/DM and the USD/CND,¹³ the new range is 3.640 to 4.881 years. This does not seem to be that much different to the range of Rogoff's, but it is nevertheless significantly different. The half-lives generated by the ESTAR are shorter and significant. Ignoring the outliers, the new range is 3.561 to 4.699. Like the results for the EQTAR, this range is significantly different to the range for the linear model even though the absolute difference appears small.

¹³ The estimate for this data series always falls out of the usual range. Part of the reason may lie on the fact that only post-unification and pre-Euro data from Germany are used. Given its small sample and its different historical background, the dynamics of the US/German exchange rate is different to the other U.S.-European pairs.

The rest of Table 4 reports the results for the Band-TAR and the MR-LSTAR. The unconditional half-life generated by either of these models is well captured by the linear estimate, as the difference of the two half-lives is insignificantly different to zero. The only two exceptions appear for the Band-TAR model—the linear half-lives for both the USD/CND and the USD/£ exchange rates are significantly different than their nonlinear counterparts at a level of 1%. Again, the results for the MR-LSTAR are the strongest.

Recently, Kiliç (2005) used a non-GIRF based computation of PPP half-life as well as the confidence intervals under a similar framework.¹⁴ His results are supportive to the exponential STAR (or ESTAR) model—a nonlinear autoregressive model with an exponential smooth transition function—when it is applied on U.S.-European bilateral real exchange rates. The half-lives generated for the non-Euro group, however, are more persistent. His estimates make use of the sample information of the threshold variable but not that of the history (as initial conditions for the impulse response simulation) and the estimated shocks with different sizes and signs. The Monte Carlo experiment in this paper makes use of all of the information. This may explain why mine is less favorable to the ESTAR model. Taylor, Peel and Sarno (2001) adopted a simulation method motivated by Gallant et al (1993) which is similar to the GIRF for their ESTAR model. In general, their simulation generates much less persistent half-life estimates, conditional on the same size of initial shocks. My Monte Carlo shares the same assumption of average initial history but

¹⁴ Kiliç (2005) uses a computation of half-life without using the GIRF; his confidence intervals for the estimate are computed from simulations. His measure is an estimator adjusted for the nonlinearities in the autoregressive coefficient, condition on the estimated value for the smooth transition function and the threshold variable: $\max(\ln 0.5 / \ln(1 + \varphi^* F(q_{t-d})), 0)$; inferences are then drawn from the distribution of the estimate. Note that he studies the exponential STAR model and thus $F(\cdot)$ denotes an exponential function. Even though the empirical value of the exponential function enters in Kiliç's nonlinear half-life equation, no impulse responses are simulated; therefore, the variation of history and shocks that are demanded by the GIRF are not implied by the formula.

not the conditionality of shocks. Hence, it will be inappropriate to compare their conditional half-life estimates with the unconditional half-life estimates reported in this paper.

5. Concluding Remarks

This paper develops a Monte Carlo experiment, incorporating the simulation technique developed by Koop, Pesaran and Potter (1997) and Potter (2000), to evaluate the half-life estimates of PPP reported in Rogoff (1996). While many nonlinear studies focus on comparing the regime-specific estimates in their threshold autoregressive or smooth transition autoregressive models and Rogoff's full-sample estimates, I used the experiment to compare the results from both linear and nonlinear models unconditional upon regimes.

Regardless of which nonlinear models is assumed as the true DGP, the unconditional half-life still falls into the range of 3-5 years. Nevertheless, the Monte Carlo simulation implies that estimations based on linear models will result in statistically different half-life if the EQTAR and the ESTAR are used. The PPP puzzle appears to be consistent with the Band-TAR and MR-LSTAR models. There are two important implications of this finding. Because the equilibrium adjustment within the transaction-cost bands is relatively slower for the Band-TAR and the MR-LSTAR than for the EQTAR, the results in the paper paint a clearer picture of the type of nonlinearities for the real exchange rate adjustment.¹⁵ In addition, although the MR-LSTAR contains more parameters than the others and thus is

¹⁵ Under the Band-TAR specification, a large PPP deviation will cause an equilibrium adjustment, which ceases in the middle regime. Under the EQTAR specification, a large PPP deviation will cause an equilibrium adjustment, which continues in the middle regime until the equilibrium is reached. As the MR-LSTAR is similar to the Band-TAR, it is reasonable to state that equilibrium adjustments are slower in the middle regime for the Band-TAR and the MR-LSTAR.

more prone to be affected by parametric uncertainty, it appears to explain the PPP puzzle the best. More attention should be paid to this relatively new model when the real exchange rate is concerned.

Reference

- Bec, F., Ben Salem, M., Carrasco, M. 2004. Detecting mean revision in real exchange rates from a multiple regime STAR model. Working Paper, Department of Economics, University of Rochester.
- Beaurdy, P., G. Koop 1993. Do recessions permanently change output? *Journal of Monetary Economics* **31**, 149-163.
- Berben, R.-P. 2000. *Macroeconomic Fundamental and Exchange Rates*. University of Amsterdam Ph.D. Thesis.
- Chen, S.-S., Engel, C. 2005. Does “aggregation bias” explain the PPP puzzle? *Pacific Economic Review* **10**, 49-72.
- Gallant, A.R., Rossi, P.E., Tauchen, G. 1993. Nonlinear dynamic structures. *Econometrica* **61**: 871-908.
- Hansen, B. 1997. Inference in TAR models. *Studies in Nonlinear Dynamics and Econometrics* **2**, Article 1.
- Koop, G., Pesaran, M.H., Potter, S.M. 1996. Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics* **74**: 119-147.
- Imbs, J., Mumtaz J., Ravn, M.O., Rey H. 2003. Nonlinearities and real exchange rate dynamics. *Journal of the European Economic Association* **1**: 639-649.
- Imbs, J., Mumtaz, J., Ravn, M.O., Rey H. 2005. PPP strikes back: aggregation and the real exchange rate. *The Quarterly Journal of Economics* **120**: 1-44.
- Kiliç, R. 2005. Confidence intervals for half-life deviations from purchasing power parity in a nonlinear model. Working Paper, Department of Economics, Georgia Institute of Technology.

- Lo, M.C., Zivot E. 2001. Threshold cointegration and nonlinear adjustment to the law of one price. *Macroeconomic Dynamics* **5**: 533-576.
- Luukkonen, R., Saikkonen P., Teräsvirta T. 1988. Testing linearity against smooth transition autoregressive models. *Biometrika* **75**: 491-499.
- Michael, P., Nobay, A.R, Peel, D.A. 1997. Transaction costs and nonlinear adjustment in real exchange rates: an empirical investigation. *Journal of Political Economy* **105**: 862-879.
- Obstfeld, M., Taylor, A.M. 1997. Nonlinear aspects of goods-market arbitrage and adjustment: Heckscher's commodity points revisited. *Journal of the Japanese and International Economies* **11**: 441-479.
- O'Connell, P.G.J., Wei, S.-J. 2002. "The bigger they are, the harder they fall" Retail price differences across U.S. cities. *Journal of International Economics* **56**, 21-53.
- Potter, S.M. 2000. Nonlinear impulse response functions. *Journal of Economic Dynamics and Control* **24**: 1425-1446.
- Sarno, L. 2003. Nonlinear exchange rate models: a selective overview. *Rivista di Politica Economica* **93**: 3-46.
- Sarantis, N. 1999. Modeling non-linearities in effective real exchange rates. *Journal of International Money and Finance* **18**: 27-45.
- Shintani, M. 2006. A non-parametric measure of convergence toward purchasing power parity. *Journal of Applied Econometrics* **21**: 589-604.
- Taylor, M.P., Sarno, L. 2003. *The Economics of Exchange Rates*. Cambridge University Press: Cambridge, U.K.

Taylor, M.P., Peel, D.A., Sarno, L. 2001. Nonlinear mean-reversion in real exchange rates: towards a solution of the purchasing power parity puzzles. *International Economic Review* **42**: 1015-1042.

Teräsvirta, T., Anderson, H. 1992. Charactering nonlinearities in business cycles using smooth transition autoregressive models. *Journal of Applied Econometrics* **7**: S119-136.

Van Dijk, D., Franses, P.H., Boswijk, H.P. 2005. Absorption of shocks in nonlinear autoregressive models, *Computational Statistics and Data Analysis*, forthcoming.

Figure 1: Regimes under a TAR Model

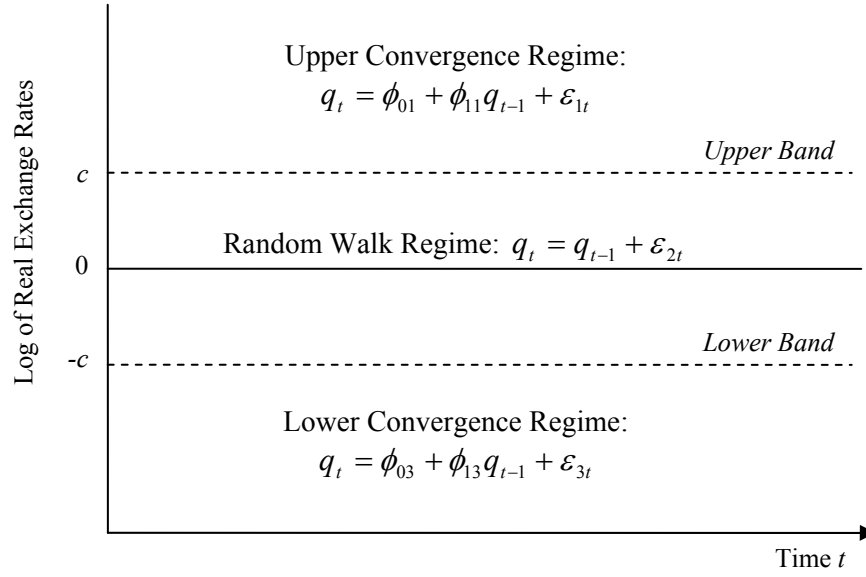
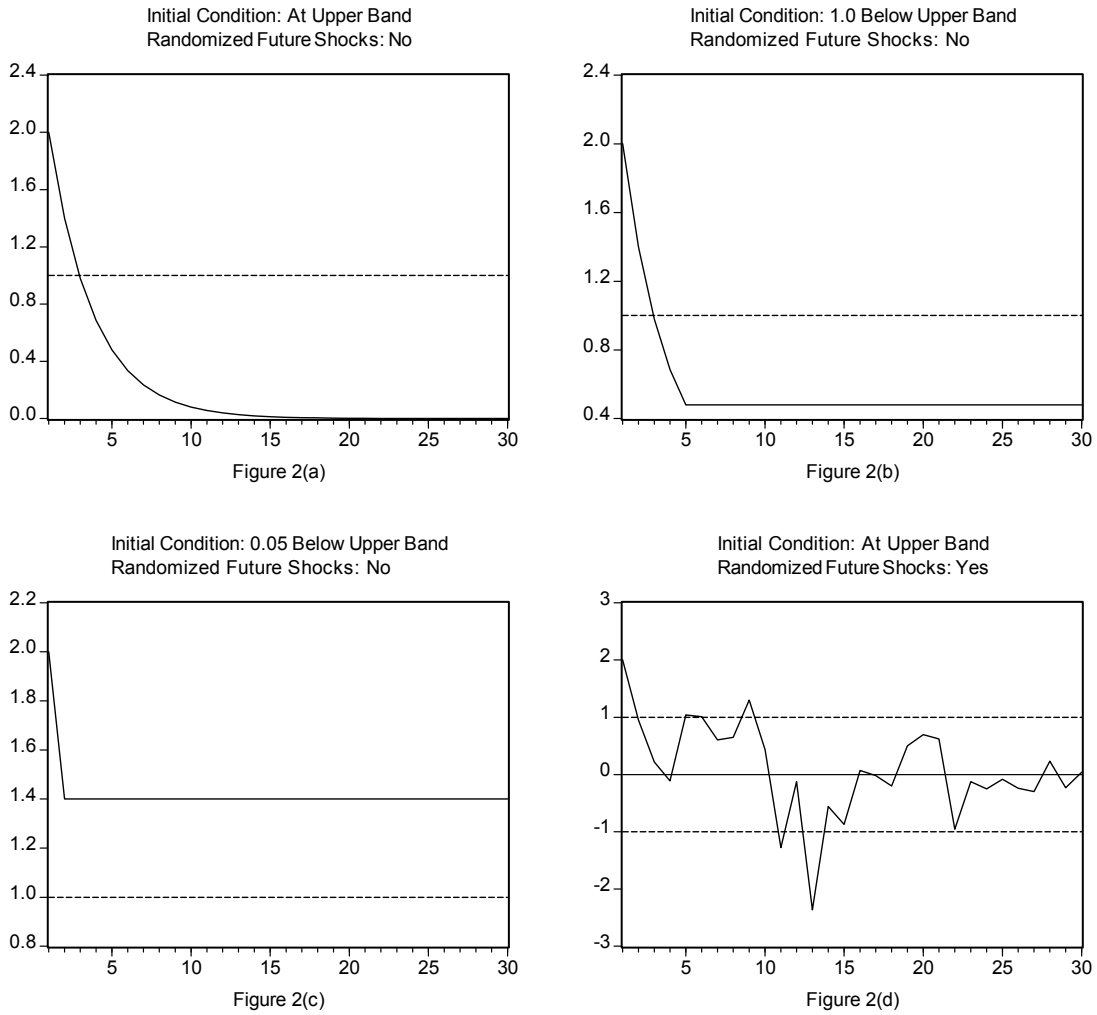


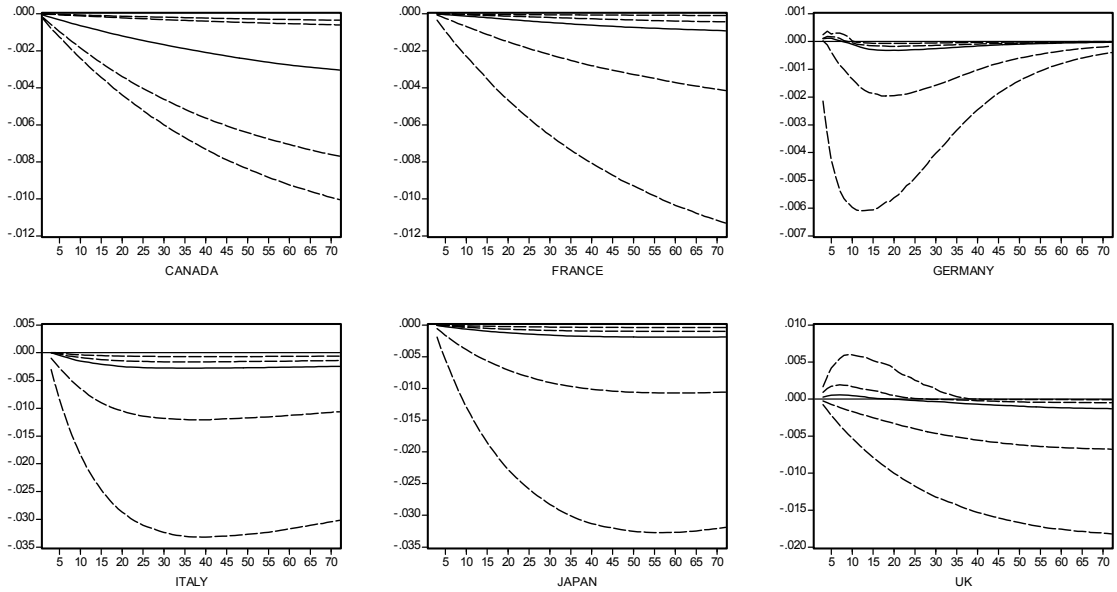
Figure 2: Impulse Response Simulation Based on an EQTAR Model



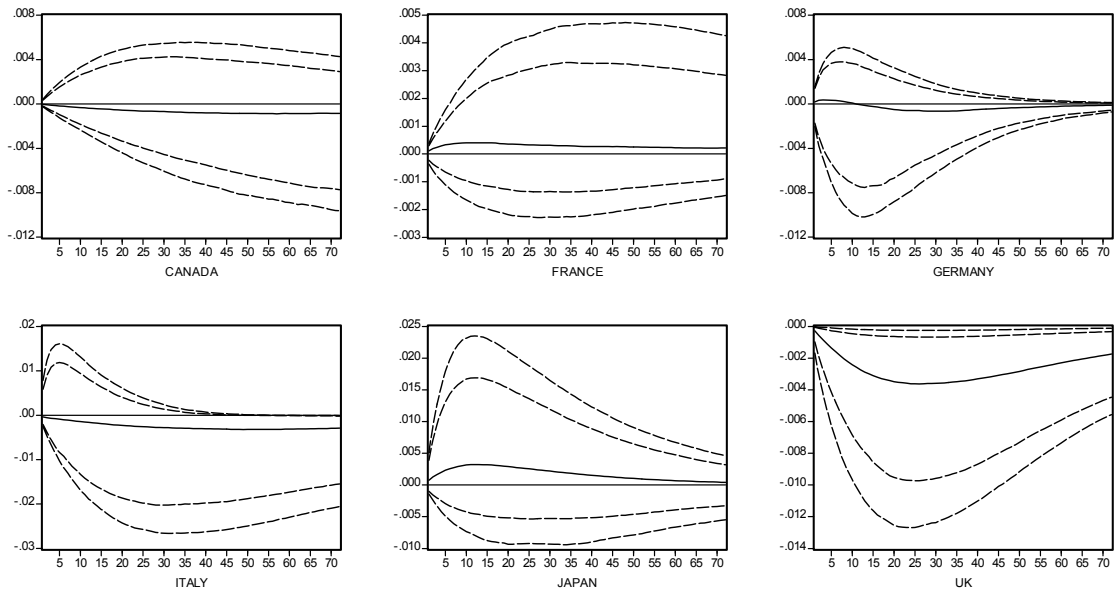
Note: The impulse response is based on equation (1) with EQTAR restrictions and $\phi_{11} = \phi_{13} = -0.3$ and $c = 1.5$. The zero line represents the initial condition. The size of the initial shock is 2 and thus the half-life can be traced from the line of 1 (and/or -1) for Figure 2(d). Future shocks for Figure 2(d) are randomized from a normal distribution with a standard deviation of 1.5 units.

Figure 3: Monte Carlo Median and Confidence Intervals of the Mean Bias

(a) EQTAR

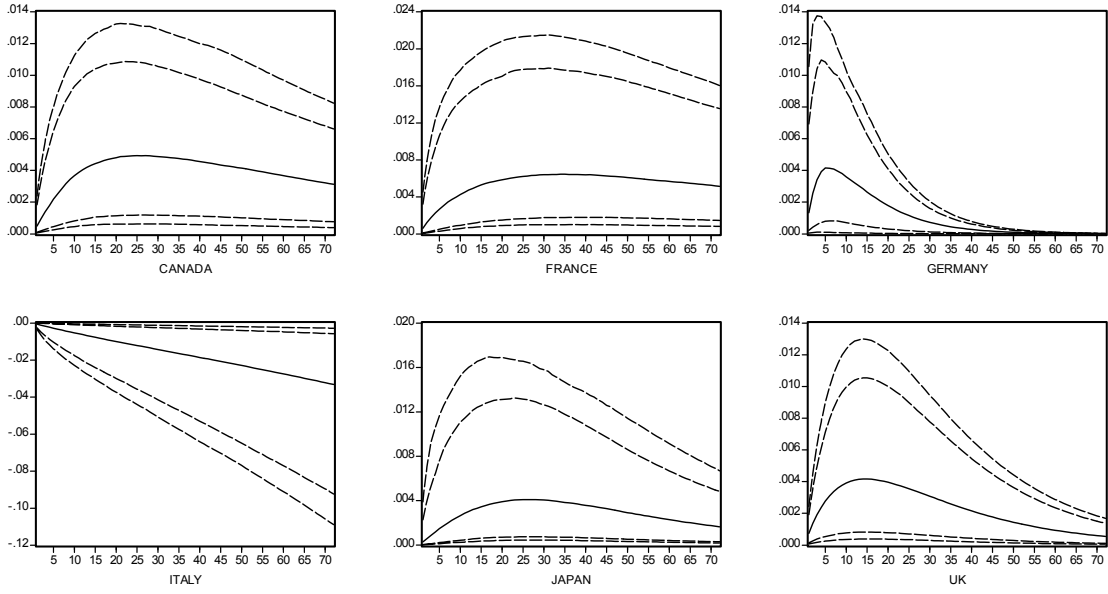


(b) Band-TAR



Note: The closest confidence bands engulf 64% of the simulated $MBI(k, \mathbf{V}; \Delta_j, \Theta_j)$ the others engulf 95%. All results were multiplied by 100 for better illustration.

(c) ESTAR



(d) MR-LSTAR

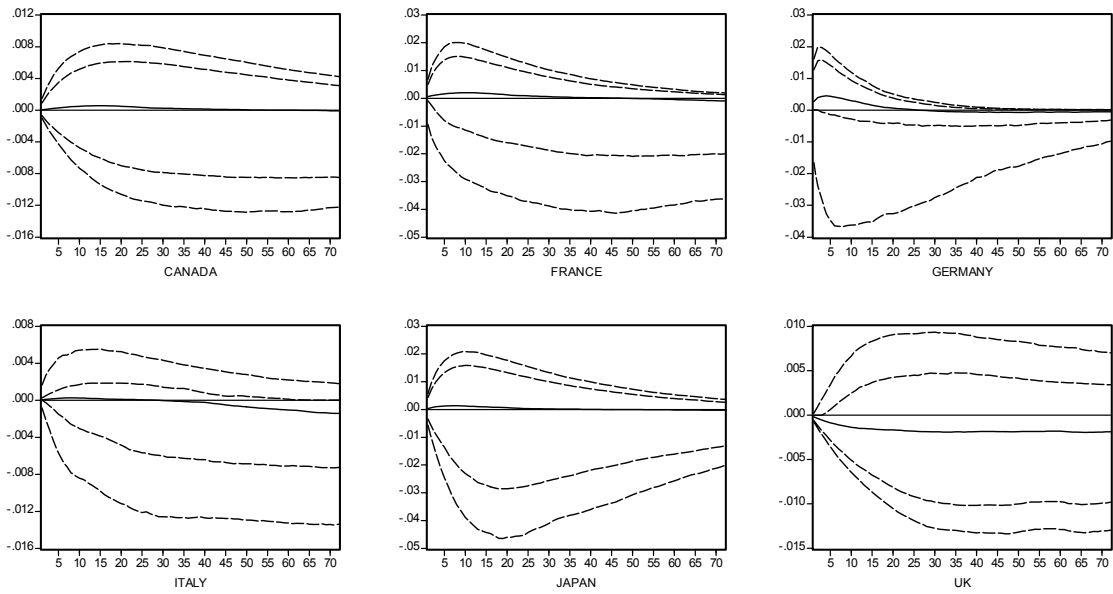


Table 1: Monte Carlo Mean Estimates of a Linear AR(1) Model with Nonlinear Data Generating Process

Panel (a): Band-TAR Model

| | | Standard Deviation Of Shocks Outside the Bands | | | | | |
|--|------------|--|---------------------|---------------------|---------------------|---------------------|---------------------|
| | | <i>0.5</i> | <i>1.0</i> | <i>1.5</i> | <i>2.0</i> | <i>2.5</i> | <i>3.0</i> |
| Standard Deviation of Shocks Inside the Bands | <i>0.5</i> | -0.0733 (0.0124) | -0.1153 (0.0171) | -0.1518 (0.0241) | -0.1813 (0.0310) | -0.2049 (0.0370) | -0.2218 (0.0413) |
| | <i>1.0</i> | -0.0976 (0.0133) | -0.1368 (0.0169) | -0.1680 (0.0216) | -0.1924 (0.0260) | -0.2111 (0.0296) | -0.2255 (0.0314) |
| | <i>1.5</i> | -0.1147 (0.0137) | -0.1491 (0.0168) | -0.1759 (0.0204) | -0.1973 (0.0238) | -0.2136 (0.0268) | -0.2267 (0.0287) |
| | <i>2.0</i> | -0.1295 (0.0139) | -0.1589 (0.0164) | -0.1825 (0.0196) | -0.2010 (0.0233) | -0.2163 (0.0256) | -0.2276 (0.0272) |
| | <i>2.5</i> | -0.1425 (0.0143) | -0.1681 (0.0164) | -0.1888 (0.0196) | -0.2051 (0.0221) | -0.2185 (0.0243) | -0.2294 (0.0264) |
| | <i>3.0</i> | -0.1543 (0.0144) | -0.1763 (0.0164) | -0.1940 (0.0190) | -0.2089 (0.0214) | -0.2209 (0.0240) | -0.2309 (0.0257) |

Panel (b): EQTAR Model

| | | Standard Deviation Of Shocks Outside the Bands | | | | | |
|--|------------|--|---------------------|---------------------|---------------------|---------------------|---------------------|
| | | <i>0.5</i> | <i>1.0</i> | <i>1.5</i> | <i>2.0</i> | <i>2.5</i> | <i>3.0</i> |
| Standard Deviation of Shocks Inside the Bands | <i>0.5</i> | -0.1452 (0.0153) | -0.1835 (0.0215) | -0.2224 (0.0314) | -0.2538 (0.0409) | -0.2753 (0.0484) | -0.2884 (0.0517) |
| | <i>1.0</i> | -0.2259 (0.0180) | -0.2500 (0.0241) | -0.2693 (0.0310) | -0.2835 (0.0361) | -0.2931 (0.0399) | -0.2983 (0.0404) |
| | <i>1.5</i> | -0.2581 (0.0176) | -0.2731 (0.0233) | -0.2835 (0.0287) | -0.2921 (0.0331) | -0.2972 (0.0357) | -0.3005 (0.0373) |
| | <i>2.0</i> | -0.2738 (0.0168) | -0.2836 (0.0216) | -0.2904 (0.0263) | -0.2954 (0.0312) | -0.2996 (0.0336) | -0.3013 (0.0349) |
| | <i>2.5</i> | -0.2828 (0.0157) | -0.2894 (0.0200) | -0.2942 (0.0250) | -0.2978 (0.0285) | -0.3006 (0.0313) | -0.3023 (0.0333) |
| | <i>3.0</i> | -0.2883 (0.0147) | -0.2931 (0.0187) | -0.2960 (0.0230) | -0.2992 (0.0268) | -0.3012 (0.0300) | -0.3027 (0.0319) |

Note: Monte Carlo standard errors are reported in parentheses.

Table 3: Estimates for the “Outside-Band” Autoregressive Coefficient in the Nonlinear Autoregressive Models

| | AR | | EQTAR | | | Band-TAR | | | ESTAR | | | MR-LSTAR | | |
|----------------|--------|------|--------|------|----------|----------|------|----------|--------|-------|----------|----------|------|----------|
| | ϕ | HL | ϕ | HL | Δ | ϕ | HL | Δ | ϕ | HL | Δ | ϕ | HL | Δ |
| Canada | -0.007 | 8.71 | -0.006 | 9.19 | +0.48 | -0.007 | 8.01 | -0.70 | -0.006 | 10.23 | +1.52 | -0.009 | 6.25 | -2.46 |
| France | -0.016 | 3.63 | -0.016 | 3.51 | -0.12 | -0.021 | 2.78 | -0.84 | -0.028 | 2.04 | -1.58 | -0.075 | 0.74 | -2.88 |
| Germany | -0.048 | 1.18 | -0.079 | 0.70 | -0.48 | -0.174 | 0.30 | -0.88 | -0.109 | 0.50 | -0.68 | -0.337 | 0.14 | -1.04 |
| Italy | -0.015 | 3.72 | -0.021 | 2.78 | -0.94 | -0.087 | 0.63 | -3.09 | -0.027 | 2.14 | -1.58 | -0.135 | 0.40 | -3.32 |
| Japan | -0.012 | 4.73 | -0.012 | 4.89 | +0.16 | -0.025 | 2.30 | -2.43 | -0.012 | 4.67 | -0.06 | -0.070 | 0.80 | -3.94 |
| U.K. | -0.016 | 3.58 | -0.029 | 1.98 | -1.60 | -0.080 | 0.69 | -2.89 | -0.035 | 1.61 | -1.97 | -0.039 | 1.44 | -2.14 |
| <i>Average</i> | -- | 4.26 | -- | 3.84 | -0.42 | -- | 2.45 | -1.80 | -- | 3.53 | -0.73 | -- | 1.63 | -2.63 |

Note: HL is short for “half-life” and Δ denotes the difference between the half-life estimated under a nonlinear model and the half-life under the linear AR model.

Table 3: Median and Confidence Intervals for the Half-Life Estimates

| | ρ_{AR} | 1% | 5% | 10% | Median ρ_{NL} | 90% | 95% | 99% |
|----------------|----------------------|-------|-------|-------|--------------------|-------|-------|-------|
| EQ-TAR | | | | | | | | |
| Canada | 8.708 ^{***} | 8.708 | 8.710 | 8.713 | 8.739 | 8.815 | 8.849 | 8.935 |
| France | 3.627 ^{***} | 3.628 | 3.630 | 3.634 | 3.660 | 3.722 | 3.745 | 3.780 |
| Germany | 1.183 ^{***} | 1.183 | 1.184 | 1.184 | 1.190 | 1.204 | 1.208 | 1.218 |
| Italy | 3.720 ^{***} | 3.728 | 3.738 | 3.750 | 3.849 | 4.083 | 4.172 | 4.305 |
| Japan | 4.733 ^{***} | 4.739 | 4.747 | 4.759 | 4.881 | 5.192 | 5.288 | 5.609 |
| United Kingdom | 3.578 ^{***} | 3.579 | 3.582 | 3.587 | 3.640 | 3.750 | 3.801 | 3.878 |
| Band-TAR | | | | | | | | |
| Canada | 8.708 ^{***} | 8.709 | 8.711 | 8.714 | 8.752 | 8.933 | 9.010 | 9.145 |
| France | 3.627 | 3.526 | 3.578 | 3.593 | 3.624 | 3.641 | 3.650 | 3.674 |
| Germany | 1.183 | 1.159 | 1.168 | 1.172 | 1.184 | 1.209 | 1.218 | 1.231 |
| Italy | 3.720 | 3.709 | 3.717 | 3.720 | 3.754 | 3.930 | 3.999 | 4.153 |
| Japan | 4.733 | 4.545 | 4.629 | 4.660 | 4.722 | 4.789 | 4.828 | 4.898 |
| United Kingdom | 3.578 ^{***} | 3.579 | 3.581 | 3.585 | 3.611 | 3.663 | 3.685 | 3.732 |
| ESTAR | | | | | | | | |
| Canada | 8.708 ^{***} | 8.386 | 8.500 | 8.526 | 8.627 | 8.687 | 8.697 | 8.707 |
| France | 3.627 ^{***} | 3.348 | 3.415 | 3.450 | 3.561 | 3.609 | 3.617 | 3.624 |
| Germany | 1.183 [*] | 1.148 | 1.155 | 1.160 | 1.173 | 1.181 | 1.183 | 1.183 |
| Italy | 3.720 ^{***} | 3.724 | 3.740 | 3.760 | 3.960 | 4.535 | 4.744 | 5.432 |
| Japan | 4.733 ^{***} | 4.555 | 4.595 | 4.632 | 4.699 | 4.727 | 4.729 | 4.732 |
| United Kingdom | 3.578 ^{***} | 3.460 | 3.516 | 3.527 | 3.559 | 3.575 | 3.577 | 3.578 |
| MR-LSTAR | | | | | | | | |
| Canada | 8.708 | 8.255 | 8.418 | 8.482 | 8.666 | 8.842 | 8.944 | 9.178 |
| France | 3.627 | 3.526 | 3.562 | 3.581 | 3.626 | 3.847 | 4.073 | 4.831 |
| Germany | 1.183 | 1.134 | 1.153 | 1.160 | 1.177 | 1.196 | 1.306 | 1.388 |
| Italy | 3.720 | 3.580 | 3.687 | 3.715 | 3.726 | 3.794 | 3.860 | 4.053 |
| Japan | 4.733 | 4.594 | 4.651 | 4.672 | 4.734 | 4.958 | 5.086 | 5.313 |
| United Kingdom | 3.578 | 3.354 | 3.490 | 3.533 | 3.598 | 3.684 | 3.718 | 3.797 |

1. *** index estimates significantly different to the mean at 1% level. ** 5% level. * 10% level.
2. Since the space allows me to report up to 3 decimal places only, some numbers may appear the same but in fact different.