University of Victoria
Department of Economics

Econometrics Working Paper EWP0703

ISSN 1485-6441

A SADDLEPOINT APPROXIMATION TO THE DISTRIBUTION OF THE HALF-LIFE ESTIMATOR IN AN AUTOREGRESSIVE MODEL: NEW INSIGHTS INTO THE PPP PUZZLE

Qian Chen

China Academy of Public Finance and Public Policy,
Central University of Finance and Economics,
Beijing, People's Republic of China

&

David E. Giles

Department of Economics, University of Victoria,
Victoria, B.C., Canada

May 2007

Abstract

We derive saddlepoint approximations for the density and distribution functions of the half-life estimated by OLS from an AR(1) or AR(p) model. Our analytic results are used to prove that none of the integer-order moments of these half-life estimators exist. This provides an explanation for the unreasonably large estimates of persistency associated with the purchasing power parity "puzzle", and it also explains the excessively wide confidence intervals reported in the empirical PPP literature.

Keywords: Saddlepoint approximation; half-life estimator; PPP puzzle

JEL Classifications: C13, C22, F31, F41

Author Contact:
David Giles, Dept. of Economics, University of Victoria, P.O. Box 1700, STN CSC, Victoria, B.C., Canada V8W 2Y2; e-mail: dgiles@uvic.ca; FAX: (250) 721-6214

A SADDLEPOINT APPROXIMATION TO THE DISTRIBUTION OF THE HALF-LIFE ESTIMATOR IN AN AUTOREGRESSIVE MODEL: NEW INSIGHTS INTO THE PPP PUZZLE

1. Introduction

Purchasing power parity (PPP) is a theory about the exchange rate between two currencies. Basically, it means that the price for a given basket of services and goods should be the same in two countries, if measured in the same currency. PPP is a building block in international economics. As PPP is a corner stone of international economics, its validity has attracted considerable interest, especially since the advent of flexible exchange rates in the early 1970's. As the real exchange rate is the nominal exchange rate adjusted for the relative price level, the tradition in the literature is to use the real exchange rate to explore PPP theory.

Essentially, there are two empirical puzzles associated with PPP. The first puzzle is the non-stationary behavior of the real exchange rate. PPP theory can be simply re-stated as saying that the real exchange rate is mean reverting. Although few economists view PPP as a short-term phenomenon, non-stationarity implies that PPP theory does not hold even in the very long run. The second empirical puzzle is that the observed degree of persistence in real exchange rates is too high to be reconciled in terms of their short-term volatility. Financial factors, such as monetary or financial shocks, cause the volatility of the exchange rate, and in the presence of price stickiness the effect of such shocks can be exaggerated. However, the high persistence of the deviations from PPP that have been observed in a vast range of empirical studies cannot be explained simply by price stickiness.

In the empirical literature the half-life is a commonly-used measure of the persistence of the deviation from PPP. This is defined as the amount of time it takes for a unit shock to dissipate by 50%. Empirical studies appear to yield a consensus of a half-life of three to five years (*e.g.*, Abuaf and Jorion, 1990; Glen 1992; Cheung and Lai, 1994) and Rogoff (1996) coined the phrase "purchasing power parity puzzle" in reference to the high persistence of the real exchange rate relative to PPP theory, even allowing for stickiness. This puzzle continues to attract considerable attention in the literature, and this provides part of the motivation for this paper. In addition, several authors have reported confidence intervals for their half-life estimates. It is also puzzling that these intervals are generally so wide as to be of no practical use.

In this paper we obtain an analytic approximation to the distribution of the half-life estimator used in the PPP literature. We show, *inter alia*, that this distribution has no finite integer-order moments. This provides an explanation for relatively large half-life estimates that have been reported, as reveals that the associated confidence intervals are spurious. The paper is constructed as follows. Section 2 reviews various PPP convergence studies that focus on half-life measures. In section 3 we derive the density and distribution functions for the half-life estimator in the AR(1) model, and explore some of its properties. These results are extended to the case of the AR(p) model in section 4; and some robustness issues are discussed in section 5. The final section discusses the implications of our results and provides suggestions for future research.

2. The PPP puzzle(s)

The empirical results relating to the first of the PPP puzzles noted above are mixed. Using standard unit-root tests, most early studies could not reject the hypothesis of a unit root in real exchange rates under floating exchange rate regimes (*e.g.*, Meese and Rogoff, 1988; Edison and Fisher, 1991; Grilli and Kaminsky, 1991). Subsequent research focused on the use of long-term historical data (*e.g.*, Diebold *et al.,* 1991; Lothian and Taylor, 1996) and the application of more powerful tests. Panel unit root tests were found of reject the unit root hypothesis, and favour PPP (*e.g*., Lothian, 1997; Wu, 1996; Papell and Theodoridis, 1998; Pedroni, 2004). The same conclusion was reached by Taylor and Sarno (1998) and Taylor *et al*. (2001), who used more powerful multivariate tests, and Chuang and Lai (1998) who used the tests suggested by Elliot *et al*. (1996) and Park and Fuller (1995). Therefore, by covering longer- term data or exploring more powerful tests, the first PPP puzzle seems to be solved: PPP theory holds, at least in the long run. Accordingly, more recent papers in this field focus on the second PPP puzzle noted above, and they explore possible reasons for over-estimating the persistence of the real exchange rate.

A simple estimator of the half-life of adjustment can be based on the linear AR(1) model

$$y_t = \alpha y_{t-1} + u_t \ ; \quad t = 0, 1, 2, \ldots, T \tag{1}$$

where y_t is the variable of interest (here, the real exchange rate), with initial value y_0, and $u_t \sim i.i.d.N(0,\sigma^2)$. The normality assumption is not needed for the construction of a half-life measure. It is used to establish our main results, but their robustness to this assumption is also discussed. The half-life for the speed of adjustment can be estimated as:

$$\hat{h} = \log(0.5)/\log(\hat{\alpha}) \quad , \tag{2}$$

where $\hat{\alpha}$ is the OLS estimator of α in (1), namely $\hat{\alpha} = (y'_{-1}y_{-1})^{-1}y'_{-1}y$, and we require $\hat{\alpha} \in (0\,,1)$ for the model to be dynamically stable, and for the estimated half-life to be positive.

The second puzzle relates to half-life estimates that are "too large" (greater than about three years, say) to be reconciled with PPP theory, with associated confidence intervals that are far too wide to provide any useful information. One possible reason for this is the bias of the OLS estimator in (1). This bias is negative in small samples, and increases with the persistence of the series. Andrews' (1993) median-unbiased estimator for AR(p) models provides a good tool to correct the bias. Unfortunately, the studies applying the median-unbiased estimator do not find support for PPP theory (*e.g*., Murray and Papell, 2002; Cashin and McDermott, 2003; Caporale *et al.*, 2005; Lopez *et al*., 2004). The results based on the median-unbiased estimator yield an estimated half-life that is higher than its OLS counterpart, and the confidence intervals are still so wide that no strong conclusions can be made about the PPP puzzle. Murray and Papell (2005) extended the median-unbiased estimation method to the panel data context, and argued that the shorter half-life of 2-2.5 year based on estimators unadjusted by the median-unbiased estimator from the previous panel data are the results of the implication of inappropriate estimation method. Murray and Papell's results are consistent with Rogoff's PPP puzzle claim. Choi *et al*. (2004) address the bias sources in estimating the half-life of PPP from panel data and found a 5.5 year of half life for 21 OECD countries from 1948-2002. In all, the bias correction seems to drive us away from PPP theory.

Other researchers have tried to resolve the puzzle by questioning the use of model (1). In the presence of transaction costs, a nonlinear representation of the real exchange rate process is more reasonable (*e.g*., Taylor, Peel and Sarno, 2001; Baum *et al.*, 2001). In nonlinear models, the mean reversion speed depends on the size of the deviation from the long-run equilibrium level: the larger are the deviations, the lower are the half-life point estimates and the narrower are the confidence intervals, and *vice versa*. So, nonlinear models might seem to provide a solution to the PPP puzzle. However, El-Gamal and Ryu (2006) find that the nonlinear Threshold Autoregression (TAR) and Exponential Smooth Threshold Autoregression (ESTAR) models exhibit the same type of decay as the AR model and in this respect add little. Chortareas and Kapetanios (2004) suggest that the second puzzle may be caused artificially by the measure of half-life that is adopted. They suggest an alternative measure, which can reduce the half-life estimate for the AR(p) model, but coincides with the measure in (2) in the case of the AR(1) model.

We take a different position from previous studies that have tried to resolve the PPP puzzle, and make two contributions. With the exception of the Bayesian analysis of Kilian and Zha (1999), previous studies have used only Monte Carlo or bootstrap simulation to investigate the distribution of the half-life estimator. Indeed, Kim *et al.* (2006, pp. 3418-3419) observe: "First, it has an unknown and possibly intractable distribution. Second, it may not possess finite sample moments since it takes extreme values as $\hat{\alpha}$ approaches one." We provide the first analytic approximations to the density and distribution functions for the usual half-life estimator.[1] Based on the density function, we then prove that the moments of the half-life estimator do not exist, and we also extend the results to the general AR(p) model. This provides an explanation for the wide confidence intervals in all of the empirical studies, and it also implies that the second PPP puzzle may arise from the use of an invalid measure of the half-life, as is suggested by Chortareas and Kapetanios (2004).

3. Saddlepoint approximations for the distribution and density functions

3.1 Background

We see from (2) that the half-life estimator is a nonlinear transformation of the OLS estimator of the coefficient in an AR(1) model. If we know the density function of $\hat{\alpha}$, then we can determine the density function of the half-life. Fortunately, various studies have considered the properties of $\hat{\alpha}$ in (1) (*e.g.*, Phillips, 1978; Lieberman, 1994a, 1994b). We use Lieberman's results to establish the properties of the half-life estimator. Lieberman (1994b) implemented a saddlepoint approximation for the density and distribution functions for the OLS estimator in the AR(1) model. Since Daniels' (1954) seminal paper, many applications have illustrated the accuracy of saddlepoint approximations for density and distribution functions in general.[2]

For equation (1), Lieberman expresses the OLS estimator $\hat{\alpha}$ as:

$$\hat{\alpha} = \frac{v'R'_\alpha C_1 R_\alpha v}{v'R'_\alpha C_2 R_\alpha v}, \qquad v \sim N(0, \sigma^2 I), \tag{3}$$

where $C_1 = \begin{bmatrix} 0 & \frac{1}{2} & \dots & 0 & 0 \\ \frac{1}{2} & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & \frac{1}{2} \\ 0 & 0 & \dots & \frac{1}{2} & 0 \end{bmatrix}_{(T+1)x(T+1)}$, $C_2 = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix}_{(T+1)x(T+1)}$,

and $R_\alpha = \begin{bmatrix} \mathrm{b} & 0\dots & & & 0 \\ \alpha \mathrm{b} & 1 & 0\dots & & 0 \\ \alpha^2 b & \alpha & 1 & 0\dots & 0 \\ \vdots & \vdots & & & \vdots \\ \alpha^T b & \alpha^{T-1} & \cdots & \alpha & 1 \end{bmatrix}_{(T+1)x(T+1)}$ $, b = \begin{cases} (1-\alpha^2)^{-1/2} & ; \text{ if } \alpha \in (\text{-}1,1) \\ 0 & ; \text{ otherwise,} \end{cases}$

Then Lieberman derived the saddlepoint approximation for the density of $\hat{\alpha}$ as

$$\hat{f}(\hat{\alpha}) = \frac{\left\{tr(\hat{A}^{-1} R'_\alpha C_2 R_\alpha)\right\} \left|\hat{A}\right|^{-\frac{1}{2}}}{\left[4\pi \, tr\{(\hat{A}^{-1}D)^2\}\right]^{\frac{1}{2}}} , \tag{4}$$

where $D = D(\hat{\alpha}) = R'_\alpha (C_1 - \hat{\alpha} C_2) R_\alpha$, $\hat{A} = A(\hat{w}) = I - 2\hat{w}D$ and $\hat{w}$ satisfies

$$tr(\hat{A}^{-1}D) = 0. \tag{5}$$

Then he approximated the distribution function of $\hat{\alpha}$ by integrating the density function and applying the Lugannani-Rice (1980) formula:

$$\hat{F}(\hat{\alpha}) = P(\hat{\alpha} < x) = \Phi(\hat{\varepsilon}) + \phi(\hat{\varepsilon})(\frac{1}{\hat{z}} - \frac{1}{\hat{\varepsilon}}) , \tag{6}$$

where $\hat{\varepsilon} = \left(\log\left|\hat{A}\right|\right)^{\frac{1}{2}} \mathrm{sgn}(\hat{w})$, $\hat{z} = \hat{w}\left[2tr\{(\hat{A}^{-1}D)^2\}\right]^{\frac{1}{2}}$, $D = D(x)$; Φ and ϕ are the standard normal distribution and density functions respectively, and $\hat{w}$ is defined by (5).

Lieberman compared the approximation of the distribution with the exact values obtained using Davies' (1973) algorithm for the c.d.f. of a weighted sum of independent chi-square variates for different sample sizes and values of α . His comparison showed that the saddlepoint approximation is excellent over the whole interval of $\hat{\alpha}$, even for very small sample sizes (Lieberman, 1994a, Table 1).

3.2 Density and distribution functions of the half-life estimator

If the real exchange rate follows an AR(1) process, then the half-life estimator is defined by (2). Taking the first PPP puzzle to be resolved, we can ignore the unit root case and apply Lieberman's result. The saddlepoint approximation for the density of the half-life estimator can be derived from that for the density of $\hat{\alpha}$ using the transformation:

$$f(\hat{h}) = f[\hat{\alpha}(\hat{h}) | 0 < \hat{\alpha} < 1]J$$

$$= \frac{f(\hat{\alpha}(\hat{h}))J}{\Pr.(0 < \hat{\alpha} < 1)}, \tag{7}$$

where $f(\hat{\alpha}(\hat{h}))$ is the density function obtained by replacing $\hat{\alpha}$ with $(0.5)^{1/\hat{h}}$ in (4); and

$$J = \frac{(0.5)^{1/\hat{h}} \ln 2}{\hat{h}^2},$$

is the Jacobian of the transformation.

$\Pr.(0 < \hat{\alpha} < 1) = \Pr.(\hat{\alpha} < 1) - \Pr.(\hat{\alpha} < 0)$ can be calculated from (6) by letting $x = 0$ and $x = 1$. We let $C = \Pr.(0 < \hat{\alpha} < 1)$, which is a constant number.

So, the saddlepoint approximation to the density function for the half-life estimator is:

$$\hat{f}(\hat{h}) = \frac{\left\{tr(\tilde{A}^{-1}R'_{\alpha}C_2R_{\alpha})\right\}\left|\tilde{A}\right|^{-1/2}}{\left[4\pi\, tr\{(\tilde{A}^{-1}\tilde{D})^2\}\right]^{1/2}} \frac{(0.5)^{1/2}\ln 2}{C\hat{h}^2} \tag{8}$$

where $\tilde{D} = D(\hat{h}) = R'_{\alpha}(C_1 - (0.5)^{1/\hat{h}}C_2)R_{\alpha}$, $\tilde{A} = A(\hat{w}) = I - 2\hat{w}\tilde{D}$ and $\hat{w}$ satisfies

$$tr(\tilde{A}^{-1}\tilde{D}) = 0.$$

Similarly, the approximation to the distribution function of the half-life estimator is:

$$\hat{F}(\hat{h}) = P(\hat{h} < x | 0 < \hat{\alpha} < 1)$$

$$= P(\frac{\log(0.5)}{\log(\hat{\alpha})} < x) \Big/ P(0 < \hat{\alpha} < 1)$$

$$= P(\hat{\alpha} < (0.5)^{1/x}) / C. \tag{9}$$

Again, (9) can be calculated easily, using equation (6).

Based on (8), we generate the numerical values for the density for different choices of α and different sample sizes. We provide some figures to compare the density of $\hat{h}$ for different values of α with the same sample size, and for different sample size with the same value of α. More specifically, Figure 1 shows the density function for α equal to 0.8, 0.9 and 0.97 and sample sizes of 30. Figure 2 shows the density of $\hat{h}$ for sample sizes of 10, 30, and 50 and α equal to 0.95. From the first figure, we can see that the density function is highly skewed to the right, and the density moves to the right and the tails become fatter as α increases. From figure 2, we can see that the location of the density also moves to the right and the tails become fatter as the sample size increases. It is clear why relatively large half-life estimates have been reported frequently in the empirical PPP literature.

Table 1: Point Estimator and Confidence Intervals of the Half-Life for Different α Values and Sample Sizes

α	T = 10		T = 30	
	Point estimator (Median)	95% Confidence Interval	Point estimator (Median)	95% Confidence Interval
0.6	1.20	[0.27, 7.12]	1.28	[0.47, 3.23]
0.7	1.58	[0.32, 12.15]	1.78	[0.63, 5.02]
0.8	2.19	[0.39, 24.49]	2.72	[0.87, 9.56]
0.9	3.39	[0.49, 60.04]	5.06	[1.27, 33.64]
0.95	4.79	[0.61, 109.52]	8.26	[1.65, 108.48]
0.97	6.07	[0.69, 155.41]	10.95	[1.88, 197.73]

Note: both the median point estimates and the confidence intervals are calculated from (9) using code written for the SHAZAM econometrics package (Whistler *et al*., 2004).

Table 1 shows the (median) point estimate and 95% confidence interval of the half-life estimator when the true data process is an AR(1) model. We see that the point estimate increases with the sample size, and when α is greater than 0.9, which is almost always the case in the empirical studies in this field, the confidence interval is very wide. As the sample size increases, the confidence interval width decreases, but it is still quite wide. The actual meaning of the half-life estimate depends on the frequency of the real data. For yearly data, the results are obviously inconsistent with the PPP theory. However, for quarterly data, the sample size is usually over 30 and the autocorrelation coefficient is quite high, so we can expect that the PPP puzzle is still there. Therefore our results are consistent with most of the other related empirical work.

3.3 Non-existence of moments

Further insights into these characteristics of the half-life estimator can be obtained by considering the moments of its distribution. Interestingly, though, we have the following result.

Theorem 1

Let the data follow a stationary AR(1) process: $y_t = \alpha y_{t-1} + u_t$, with $u_t \sim N(0, \sigma^2)$. The half-life estimator is defined as $\hat{h} = \log(0.5)/\log(\hat{\alpha})$, where $\hat{\alpha}$ is the least squares estimator of α, and $\hat{\alpha} \in (0,1)$. Then the mean of the half-life estimator does not exist.

Proof

$$M(\hat{h}) = \int_0^\infty \hat{h}\, f(\hat{h})\, d\hat{h}\,,$$

$$= \int_0^1 \frac{\log(0.5)}{\log(\hat{\alpha})} f(\hat{\alpha})\, d\hat{\alpha}$$

Let $u(\hat{\alpha}) = \left\{tr(\hat{A}^{-1}R'_\alpha C_2 R_\alpha)\right\}\left|\hat{A}\right|^{-\frac{1}{2}}$ and $v(\hat{\alpha}) = \left[4\pi\, tr\{(\hat{A}^{-1}D)^2\}\right]^{\frac{1}{2}}$

$$\hat{M}\,(\,\hat{h}\,) = \int_0^1 \frac{\log(0.5)}{\log(\hat{\alpha})} \frac{u(\hat{\alpha})}{v(\hat{\alpha})}\, d\hat{\alpha}$$

$$= \underset{\varepsilon \to 0}{Lim} \int_\varepsilon^{1-\varepsilon} \frac{\log(0.5)}{\log(\hat{\alpha})} \frac{u(\hat{\alpha})}{v(\hat{\alpha})}\, d\hat{\alpha}\,.$$

Since the whole interval of $\hat{\alpha}$ is $(-\infty,\ \infty)$, $u(\hat{\alpha})$ and $v(\hat{\alpha})$ are continuous functions of $\hat{\alpha}$ on the closed interval $[0,\ 1]$. According to the extreme value theorem, we can assume:

(i) when $\hat{\alpha} = \overline{\alpha}$, $u(\hat{\alpha})$ gets to its minimum value N and $N \neq 0$.

(ii) when $\hat{\alpha} = \breve{\alpha}$, $v(\hat{\alpha})$ gets to its maximum value M and $M \neq \infty$.

(The justification for assumptions (i) and (ii) is given in the Appendix.)

Given that $f(\hat{\alpha}) \geq \delta$ for some $\delta > 0$ in (0, 1), then:

$$\hat{M}\,(\hat{h}) > \underset{\varepsilon \to 0}{Lim} \int_\varepsilon^{1-\varepsilon} \frac{\log(0.5)}{\log(\hat{\alpha})} \frac{N}{M}\, d\hat{\alpha}$$

$$= \log(0.5)\frac{N}{M}\underset{\varepsilon\to 0}{Lim}\int_{\varepsilon}^{1-\varepsilon}\frac{1}{\log(\hat{\alpha})}d\hat{\alpha}$$

$$= \log(0.5)\frac{N}{M}\left[\underset{\varepsilon\to 0}{Lim}(\frac{\hat{\alpha}}{\log(\hat{\alpha})}\Big|_{\varepsilon}^{1-\varepsilon}) + \underset{\varepsilon\to 0}{Lim}\int_{\varepsilon}^{1-\varepsilon}\frac{1}{[\log(\hat{\alpha})]^2}d\hat{\alpha}\right]$$

$$= \log(0.5)\frac{N}{M}\left[\infty - \underset{\varepsilon\to 0}{Lim}\int_{\varepsilon}^{1-\varepsilon}\frac{1}{[\log(\hat{\alpha})]^2}d\hat{\alpha}\right] .$$

So, the estimated mean of the half-life estimator does not exist. Based on the inversion formula, we know that

$$f(\hat{\alpha}) = \hat{f}(\hat{\alpha})(1+\ldots) \quad .$$

Therefore if the estimated mean $\hat{M}(\hat{h})$ based on the saddlepoint approximation does not exist, then the true mean $M(\hat{h})$ does not exist, either.[3]

Corollary 1

Let the data follow a stationary AR(1) process: $y_t = \alpha y_{t-1} + u_t$, with $u_t \sim N(0,\sigma^2)$. The half-life estimator is defined as $\hat{h} = \log(0.5)/\log(\hat{\alpha})$, where $\hat{\alpha}$ is the least squares estimator and $\hat{\alpha} \in (0,1)$. Then the integer-order moments of the half-life estimator do not exist.

The proof follows that of Theorem 1.

4. Properties of the half-life estimator in the AR(p) model

Some PPP studies have been based on the AR(p) model, to take account of more general features of the data. So, it is of interest to see if our results also hold in this case. In order to make the problem workable, we make some reasonable simplifying assumptions. First, we need to know the formula used to estimate the half-life in the case of the AR(p) model. Essentially there are two ways that are used to estimate the half-life for the AR(p) model in this literature. First, some studies use the impulse response function to estimate the half-life by using some nonparametric method, such as the bootstrap or Monte Carlo method. Second, other studies estimate the half-life based on the formula constructed from the coefficient estimator from an augmented Dickey-Fuller (ADF) regression equation. The empirical work based on both of these methods has found similar results for the PPP puzzle, namely an implausibly large half-life estimate and a very wide confidence interval. In order to derive a specific density and distribution functions for half-life, we use the second approach here. The basic idea is as follows.

If the real exchange rate y_t follows an AR(p) process, then:

$$y_t = \sum_{i=1}^{p} \alpha_i y_{t-i} + u_t . \tag{10}$$

There is no explicit half-life function for the AR(p) model based on the estimator of the coefficients in (10). The formula often used in practice involves approximating the half-life by estimating an ADF equation:

$$\Delta y_t = \beta y_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta y_{t-i} + u_t , \qquad u_t \sim i.i.d.N(0,\sigma^2) . \tag{11}$$

We suppose that the data are stationary, so that $\beta \in (-1,1)$. Then, based on (11), we estimate the half-life using:

$$\hat{h} = \log(0.5)/\log(1+\hat{\beta}) , \qquad\qquad 1+\hat{\beta} \in (0,1) . \tag{12}$$

In order to express the OLS estimator $\hat{\beta}$ simply, we first apply some transformations to the data. Let,

$$R_1 = M'\Delta y_t \qquad \text{and} \qquad R_2 = M' y_{t-1}$$

where $M = I - Y(Y'Y)^{-1}Y'$, and $Y = (\Delta y_{t-1} \;\; \Delta y_{t-2} \;\; \Delta y_{t-3} \cdots \; \Delta y_{t-p+1})$, and we are implicitly conditioning on the p initial observations. Then, using standard partitioning results,

$$\hat{\beta} = (R_2{}' R_2)^{-1} R_2{}' R_1 ,$$

which can be rewritten as:

$$\hat{\beta} = \frac{R'QR}{R'GR} , \tag{13}$$

where

$$Q = \begin{bmatrix} 0 & 0\ldots & & 0 & 0 & \frac{1}{2} & 0 & 0\cdots & 0 \\ 0 & 0\ldots & & 0 & 0 & 0 & \frac{1}{2} & 0\ldots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & & & \vdots \\ 0 & 0\ldots & & 0 & 0\ldots & & & & \frac{1}{2} \\ \frac{1}{2} & 0\ldots & & 0 & 0\ldots & & & & 0 \\ 0 & \frac{1}{2} & 0\ldots & 0 & 0\ldots & & & & 0 \\ \vdots & \vdots & & \vdots & \vdots & & & & \vdots \\ 0 & 0\ldots & & 0 & \frac{1}{2} & 0\ldots & & & 0 \end{bmatrix}_{2T\times 2T}$$

$$G = \begin{bmatrix} 1 & 0\ldots & & 0\ldots & & 0 \\ 0 & 1 & 0\ldots & 0\ldots & & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0\cdots & & 1 & 0\ldots & 0 \\ 0 & 0\ldots & & & & 0 \\ 0 & 0\ldots & & & & 0 \\ \vdots & \vdots & & & & \vdots \\ 0 & 0\ldots & & & & 0 \end{bmatrix}_{2T\times 2T} \quad \text{and} \quad R = \begin{bmatrix} R_2 \\ R_1 \end{bmatrix}.$$

Now, we define the covariance matrix of R to be $\Omega_{2T\times 2T}$ and $\Omega^{-1} = P'P$. Equation (13) can be written as:

$$\hat{\beta} = \frac{v'P'QPv}{v'P'GPv}, \qquad v \sim N(0, \sigma^2 I). \tag{14}$$

We can see that equations (14) and (3) are quite similar, so based on Lieberman's method, we can derive the approximation to the density function of $\hat{\beta}$ as,

$$\hat{f}(\hat{\beta}) = \frac{\left\{tr(\hat{N}^{-1}P'GP)\right\}\left|\hat{N}\right|^{-\frac{1}{2}}}{\left[4\pi\, tr\{(\hat{N}^{-1}L)^2\}\right]^{\frac{1}{2}}}, \tag{15}$$

where $L = L(\hat{\beta}) = P'(Q - \hat{\beta}G)P$, $\hat{N} = N(\hat{w}) = I - 2\hat{w}L$ and $\hat{w}$ satisfies

$$tr(\hat{N}^{-1}L) = 0 \quad . \tag{16}$$

Then, applying the Lugannani-Rice formula, the approximate distribution function of $\hat{\beta}$ is

$$\hat{F}(\hat{\beta}) = P(\hat{\beta} < x) = \Phi(\hat{\varepsilon}) + \phi(\hat{\varepsilon})(\frac{1}{\hat{z}} - \frac{1}{\hat{\varepsilon}}) \quad , \tag{17}$$

where $\hat{\varepsilon} = \left(\log\left|\hat{N}\right|\right)^{\frac{1}{2}} \operatorname{sgn}(\hat{w})$, $\hat{z} = \hat{w}\left[2tr\{(\hat{N}^{-1}L)^2\}\right]^{\frac{1}{2}}$,

$L = L(x)$;as before, Φ and ϕ are the standard normal distribution and density functions respectively, and $\hat{w}$ is defined by (16).

Let $\tilde{\alpha} = 1 + \hat{\beta}$, and using the fact that the Jacobian is unity, the approximate density function of $\tilde{\alpha}$ is :

$$\hat{f}(\tilde{\alpha}) = \frac{\left\{tr(\tilde{N}^{-1}P'GP)\right\}\left|\tilde{N}\right|^{-\frac{1}{2}}}{\left[4\pi\, tr\{(\tilde{N}^{-1}\tilde{L})^2\}\right]^{\frac{1}{2}}}, \tag{18}$$

where $\tilde{L} = L(\tilde{\alpha}) = P'(Q-(\tilde{\alpha}-1)G)P$, $\tilde{N} = N(\hat{w}) = I - 2\hat{w}\tilde{L}$ and $\hat{w}$ satisfies

$$tr(\tilde{N}^{-1}\tilde{L})=0.$$

Using (12), and allowing for the Jacobian, the density function of the half-life estimator in the AR(p) model is:

$$\begin{aligned} f(\hat{h}) &= f(\tilde{\alpha}(\hat{h}) \mid 0<\tilde{\alpha}<1)J \\ &= \frac{f(\tilde{\alpha}(\hat{h}))J}{P(-1<\hat{\beta}<0)}, \end{aligned} \tag{19}$$

where $f(\tilde{\alpha}(\hat{h}))$ is the density function by replacing $\tilde{\alpha}$ with $(0.5)^{1/\hat{h}}$ in (18); and the Jacobian is

$$J = \frac{(0.5)^{1/\hat{h}}\ln 2}{\hat{h}^2}.$$

$\Pr.(-1<\hat{\beta}<0) = \Pr.(\hat{\beta}<0) - P(\hat{\beta}<-1)$ can be calculated from (17) by letting $x=0$ and $x=1$. We let $K=\Pr.(-1<\hat{\beta}<0)$, which is a constant number.

So, the saddlepoint approximation for the density function for the half-life estimator is:

$$\hat{f}(\hat{h}) = \frac{\left\{tr(\overline{N}^{-1}P'GP)\right\}\left|\overline{N}\right|^{-\frac{1}{2}}}{\left[4\pi\, tr\{(\overline{N}^{-1}\overline{L})^2\}\right]^{\frac{1}{2}}}\frac{(0.5)^{1/\hat{h}}\ln 2}{K\hat{h}^2}, \tag{20}$$

where $\overline{L} = L(\hat{h}) = P'(Q-((0.5)^{1/\hat{h}}-1)G)P$, $\overline{N} = N(\hat{w}) = I - 2\hat{w}\overline{L}$ and $\hat{w}$ satisfies

$$tr(\overline{N}^{-1}\overline{L}) = 0.$$

Based on the density function (20), we have the following result.

Theorem 2

Suppose that the data follow a stationary AR(p) process and satisfy the ADF equation: $\Delta y_t = \beta y_{t-1} + \sum_{i=1}^{p-1}\phi_i \Delta y_{t-i} + u_t$, with $u_t \sim N(0,\sigma^2)$ and $\beta \in (-1,1)$, and the half-life is defined as $\hat{h} = \log(0.5)/\log(1+\hat{\beta})$, where $\hat{\beta}$ is the least squares estimator and $\hat{\beta} \in (-1,0)$. Then the mean of the half-life estimator does not exist.

The proof follows that of Theorem 1.

Corollary 2

Suppose that the data follow a stationary AR(p) process and satisfy the ADF equation: $\Delta y_t = \beta y_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta y_{t-i} + u_t$, with $u_t \sim N(0, \sigma^2)$ and $\beta \in (-1,1)$, and the half-life is defined as $\hat{h} = \log(0.5)/\log(1+\hat{\beta})$, where $\hat{\beta}$ is the least squares estimator and $\hat{\beta} \in (-1, 0)$. Then none of the integer-order moments of the half-life estimator exist.

The proof follows that of Theorem 1

5. Robustness results

In the previous section we assume that the data are normally distributed. Here, we consider the robustness of our results to a relaxation of this assumption. Let H_k represent the regular regression model; let H_k^{-1} represent the first-order autoregressive model; and let $E_1(n, \Sigma)$ represent the elliptically symmetric family of distributions.[4] Then King (1979; p. 121) proves that "when the disturbance vector of H_k and H_k^{-1} takes an $E_1(n, \Sigma)$ distribution, any linear unbiased or any well-behaved non-linear estimator will have very similar properties to those of the same estimator when the disturbance term is normally distributed." From King's result, we would anticipate that our own results will be robust to departures from normality, within the elliptically symmetric family of distributions.

Of primary concern here is whether the non-existence of the moments of the half-life estimator still holds under other distributional assumptions. We can apply further results of Lieberman (1997) to establish the robustness of the theorems presented above to the distributional assumption. Lieberman derives the saddlepoint approximation for the density and cumulative distribution function for the estimator $\hat{\alpha}$ in an AR(1) model with exogeneous variables. Applying his result to (3), we can get the saddlepoint approximation to the density of $\hat{\alpha}$ in (3). First, we let

$$S = v'R_\alpha' C_1 R_\alpha v - \alpha v'R_\alpha' C_2 R_\alpha v$$

$$Z = v'R_\alpha' C_2 R_\alpha v$$

$$B = R'_\alpha C_2 R_\alpha \, .$$

Then the saddlepoint approximation to the density of $\hat{\alpha}$ is:

$$\hat{f}(\hat{\alpha}) = \frac{\tilde{k}_{10} e^{\tilde{K}_S}}{\sqrt{2\pi \, \tilde{k}_2^S}} \, , \tag{21}$$

with the saddlepoint $\hat{w}$ satisfying

$$K'_S(\hat{w}) = 0 \, , \tag{22}$$

where $K_S(w)$ is the cumulant generating function of S and

$$\tilde{K}_S = K_S(\hat{w}) \tag{23}$$

$$\tilde{k}_2^S = K''_D(\hat{w}) \tag{24}$$

$$k_{10} = E(Z) \tag{25}$$

$$\tilde{k}_{10} = k_{10}(\hat{w}) \, . \tag{26}$$

Suppose v has arbitrary cumulants $k^i = 0$, $k^{i,j}$, $k^{i,j,k}, \cdots$, where the cumulants are defined as follows:

$$k^{i,j} = cum(v^i, v^j)$$

$$k^{i,j,k} = cum(v^i, v^j, v^k) \, .$$

Then (24) and (26) can be expressed in terms of v's cumulants $k^{i,j}$, $k^{i,j,k}, \cdots$.

$$\tilde{k}_2^S = \sum\nolimits_{ijkl} s_{ij} s_{kl} k^{ij,kl} \tag{27}$$

$$\tilde{k}_{10} = \sum\nolimits_{ij} b_{ij} k^{ij} \qquad . \tag{28}$$

This specification allows the v's to be correlated. When v is $i.i.d.$, (27) and (28) reduce to

$$\tilde{k}_2^S = k_4 \sum\nolimits_{ij} s_{ij}^2 + 2k_2^2 \sum\nolimits_{ij} s_{ij}^2 \tag{29}$$

$$\tilde{k}_{10} = k_2 \sum\nolimits_i b_{ii} \tag{30}$$

where $k_2 = k^{i,i}, k_4 = k^{i,i,i,i}$.

The approximating function in (21) is continuous on a closed interval $\hat{\alpha} \in [0,1]$. We can use the same procedure as for Theorem 1 to prove that the moments of the half-life estimator do not exist. Let

$$M(\hat{h}) = \int_0^{\infty} \hat{h}\, f(\hat{h})\, d\hat{h}$$

so
$$\hat{M}(\hat{h}) = \int_0^1 \frac{\log(0.5)}{\log(\hat{\alpha})} \frac{\tilde{k}_{10} e^{\tilde{K}_S}}{\sqrt{2\pi\, \tilde{k}_2^S}}\, d\hat{\alpha}\,.$$

If v is $i.i.d.$ and the second cumulant of v is finite, then $\tilde{k}_2^S$ and $\tilde{k}_{10}$ are defined by (29) and (30). And we can also see that both $\tilde{k}_2^S$ and $\tilde{k}_{10}$ are continuous functions of $\hat{\alpha}$ on the closed interval $[0,\ 1]$, and they are the sum of a finite number of terms. Therefore, there is a non-zero minimum and maximum for the numerator and denominator of the expression for the density function of $\hat{\alpha}$ in (21). We assume that $\hat{N}$ is the minimum value of the numerator and $\hat{M}$ is the maximum value of the denominator, and $\hat{N} \neq 0$, $\hat{M} \neq 0$. Then, as was the case in Theorem 1, it is readily seen that none of the integer-order moments of $\hat{h}$ exist.

Therefore, all of our main results hold as long as v is $i.i.d.$ and the second cumulant of v is finite. In addition to the normal distribution, there are many distributions with a finite second cumulant. When we allow the disturbances to be correlated, the situation is more complicated. However, we can still find quite a large class of distributions which will satisfy the conditions of the above proof. For the AR(p) model, we can apply (21) to (18). The situation is almost the same as for the AR(1) model. Therefore, the property that the moments of the half-life estimator do not exist is quite robust to the distributional assumption.

6. Conclusions

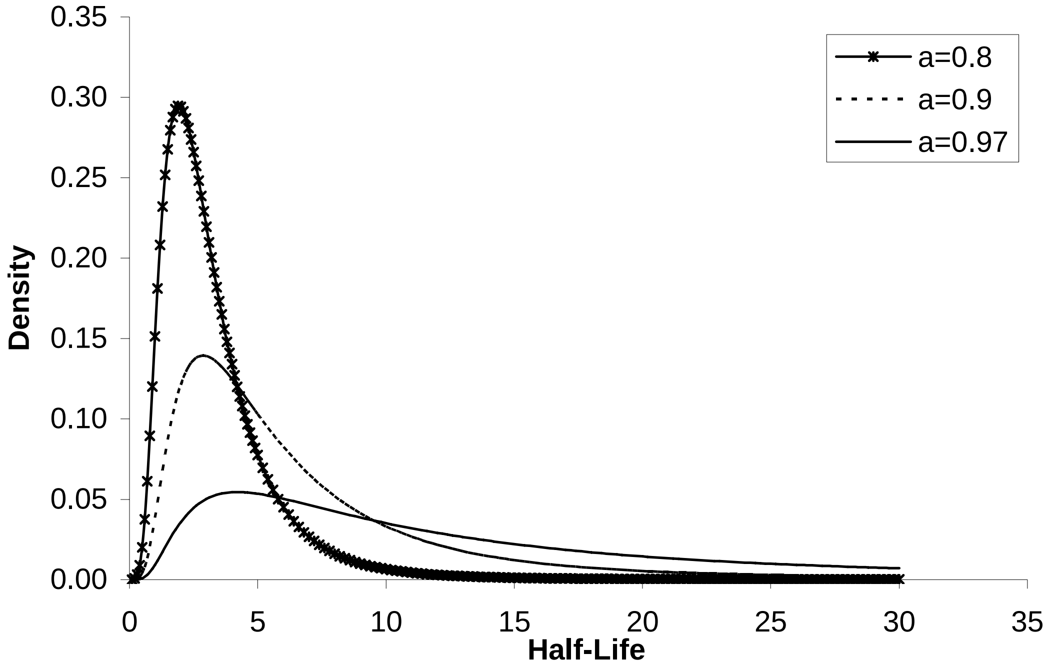
Given the important role of PPP theory in economics, it is natural that the "PPP puzzle" has attracted a lot of attention. This paper provides saddlepoint approximations for the density and distribution functions for the half-life estimators based on the OLS estimation of AR(1) or AR(p) models, and proves analytically that the moments of such half-life estimator do not exist. These result are also shown to be quite robust to the underlying distributional assumptions. These properties of the conventional half-life estimators explain both the unreasonably large point estimates, and very wide confidence intervals that have been reported in the associated empirical studies.

Our results have some implications for future research. First, the poor properties of the half-life estimator may suggest that the measure that has been traditionally used is not a good one. This is consistent with Chortareas and Kapetanios' (2004) arguments that the puzzle may be caused artificially by the measure we use. Future work may be better to focus on constructing more appropriate measures of persistence, rather than just explore all possible reasons to improve the accuracy based on the current measure of the half-life. Second, we have not considered the case of nonlinear models. However, the assumption in the nonlinear models employed in the PPP literature is that the arbitrage happens only when the deviation is quite large. So we can imagine that when the deviation is small, the situation for the nonlinear models would be similar with the case we have analyzed here. Therefore, nonlinear models can never solve the problem for small deviations. [5]

Acknowledgements

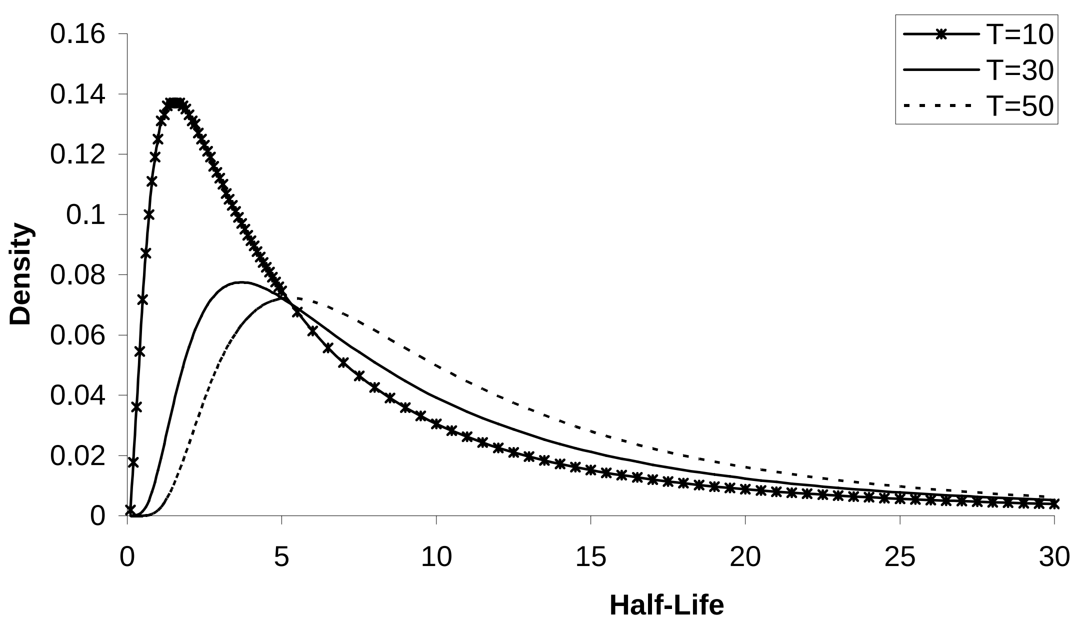
We are grateful to Judith Clarke, Nilanjana Roy, Min Tsao, Graham Voss and Aman Ullah for their very helpful comments and suggestions.

Figure 1: The density of half-life estimator when T=30



Note: Figure 1 depicts the saddlepoint density function (8), which is the approximate density of $\hat{h}$ when the sample size is 30 and the true α is 0.8, 0.9, 0.97 respectively.

Figure 2: The density of half-life estimator when a=0.95



Note: Figure 2 depicts the saddlepoint density function (8), which is the approximate density of $\hat{h}$ when the sample size is 10, 30, 50 respectively and the true α is 0.95.

Appendix: Justification of assumptions that $N \neq 0$ and $M \neq \infty$

Here we justify assumptions (i) and (ii) used in the proof of Theorem 1.

First, we prove that $N \neq 0$:

$$N = u(\overline{\alpha}) = \left\{tr(\hat{A}^{-1}R'_\alpha C_2 R_\alpha)\right\}\left|\hat{A}\right|^{-\frac{1}{2}} .$$

So, if N =0, then $\left|\hat{A}\right|^{-\frac{1}{2}} = 0$ or $\left\{tr(\hat{A}^{-1}R'_\alpha C_2 R_\alpha)\right\} = 0$.

As the density exists, we can rule out the possibility that $\left|\hat{A}\right|^{-\frac{1}{2}} = 0$.

For
$$\left\{tr(\hat{A}^{-1}R'_\alpha C_2 R_\alpha)\right\} = \sum_{t=0}^{T} \frac{f_t}{1-2\hat{w}d_t} ,$$

where the d_t are the eigenvalues of matrix D and the f_t are the eigenvalues of $R'_\alpha C_2 R_\alpha$

Since $\left|\hat{A}\right|^{-\frac{1}{2}} = \exp\left\{-\frac{1}{2}\sum_0^T \log(1-2\hat{w}d_i)\right\}$ exists, $\frac{1}{1-2\hat{w}d_t}$ must be positive. Also, $R'_\alpha C_2 R_\alpha$ is a positive definite matrix, so the eigenvalues f_t are all positive. Therefore:

$$\left\{tr(\hat{A}^{-1}R'_\alpha C_2 R_\alpha)\right\} > 0 ,$$

and so $N \neq 0$.

Second, we prove that $M \neq \infty$.

$$M = v(\tilde{\alpha}) = \left[4\pi tr\{(\hat{A}^{-1}D)^2\}\right]^{\frac{1}{2}} = \left\{4\pi \sum_0^T \left[d_t^2 / (1-2\hat{w}d_t)^2\right]\right\}^{1/2}$$

so, if $M = \infty$, it must be the case that $\frac{1}{(1-2\hat{w}d_t)^2}$ is zero. But from $\left|\hat{A}\right|^{-\frac{1}{2}} \neq 0$, we know this cannot hold. So, $M \neq \infty$.

References

Abuaf, N., Jorion, P., 1990. Purchasing power parity in the long run. Journal of Finance 45, 157-174.

Andrews, D. W. K., 1993. Exactly median-Uubiased estimation of first order autoregressive/unit root models. Econometrica 61, 139-165.

Baum, C. F., Barkoulas, J. T., Caglayan, M., 2001. Nonlinear adjustment to purchasing power parity in the post-Bretton Woods era. Journal of International Money and Finance 20, 379-399.

Caporale, G. M., Cerrato, M., Spagnolo, N., 2005. Measuring half-lives using a non-parametric bootstrap approach. Applied Financial Economics Letters 1, 1-4.

Cashin, P., McDermott, C. J., 2003. An unbiased appraisal of purchasing power parity. International Monetary Fund Staff Papers 50, 321-351.

Cheung, Y. W., Lai, K. S., 1994. Mean reversion in real exchange rates. Economics Letters 46, 251-256.

Cheung, Y. W., Lai, K. S., 1998. Parity reversion in real exchange rates during the post-Bretton Woods period. Journal of International Money and Finance 17, 97-614.

Cheung, Y. W., Lai, K. S., Bergman, M., 2004. Dissecting the PPP puzzle: The unconventional roles of nominal exchange rate and price adjustments. Journal of International Economics 64, 135-150.

Chmielewski. M. A., 1981. Elliptically symmetric distributions: A review and bibliography. International Statistical Review 49, 67-74.

Choi, C. Y., Mark, N. C., Sul. D., 2004. Unbiased estimation of the half-life to PPP convergence in panel data. National Bureau of Economic Research Working Paper No. W10641.

Chortareas, G., Kapetanios, G., 2004. How puzzling is the PPP puzzle? An alternative half-life measure of convergence to PPP. University of London, Working Paper No. 522.

Daniels, H. E., 1954. Saddlepoint approximations in statistics. Annals of Mathematical Statistics 25, 631-650.

Davies, R. B., 1973. Numerical inversion of a characteristic function. Biometrika, 60, 415-417.

Diebold, F. X., Husted, S., Rush, M., 1991. Real exchange rates under the gold standard. Journal of Political Economy 99, 1252-1271.

Edison, H. J., Fisher, E. O., 1991. A long-run view of the European monetary system. Journal of International Money and Finance 10, 53-70.

El-Gamal, M.A., Ryu, D., 2006. Short-memory and the PPP hypothesis. Journal of Economic Dynamics and Control 30, 361-391.

Elliott, G., Rothenberg, T. J., Stock, J. H., 1996. Efficient tests for an autoregressive unit root. Econometrica 64, 813-836.

Engel, C., Morley, J. C., 2001. The adjustment of prices and the adjustment of the exchange rate. NBER Working Paper No.8550.

Giles, D. E. A., 2001. A saddlepoint approximation to the distribution function of the Anderson-Darling test statistic. Communications in Statistics, B 30, 899-905.

Glen, J. H., 1992. Real exchange rates in the short, medium, and long run. Journal of International Economics 33, 147-166.

Goustis, C., Casella, G., 1999. Explaining the saddlepoint approximation. American Statistician 53, 216-224.

Grilli, V., Kaminsky, Z. G., 1991. Nominal exchange rate regimes and the real exchange rate. Journal of Monetary Economics 27, 191-212.

Huzurbazar, S., 1999. Practical saddlepoint approximations. American Statistician 53, 225-232.

Kariya, T., Eaton, M. L., 1977. Robust tests for spherical symmetry. Annals of Statistics 5, 206-215.

Kilian, L., Zha, T., 1999. Quantifying the half-life of deviations from PPP: The role of economic priors. Federal Reserve Bank of Atlanta Working Paper 99-21.

Kim, J. H., Silvapulle, P., Hyndman, R. J., 2006. Half-life estimation based on the bias-corrected bootstrap: A highest density region approach. Computational Statistics and Data Analysis 51, 3418-3432.

King, M. L., 1979. Some aspects of statistical inference in the linear regression model. Ph.D. dissertation, University of Canterbury, New Zealand.

King, M. L., 1980. Robust tests for spherical symmetry and their application to least squares regression. Annals of Statistics 8, 1265-1271.

Lieberman, O., 1994a. Saddlepoint approximation for the distribution of a ratio of quadratic forms in normal variables. Journal of the American Statistical Association 89, 924-928.

Lieberman, O., 1994b. Saddlepoint approximation for the least squares estimator in first-order autoregression. Biometrika 81, 807-811.

Lieberman, O., 1997. The effect of nonnormality. Econometric Theory 13, 52-78.

Lopez, C., Murray, C. J., Papell, D. H., 2004. More powerful unit root tests and the PPP puzzle. Working paper, Department of Economics, University of Houston.

Lothian, J. R., 1997. Multi-country evidence on the behavior of purchasing power parity under the current float. Journal of International Money and Finance 16, 19-35.

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

Lugannani, R., Rice, S. O., 1980. Saddlepoint approximation for the distribution of the sum of independent random variables. Advances in Applied Probability 12, 475-90.

Meese, R. A., Rogoff, K. S., 1988. Was it real? The exchange-rate interest differential relation over the modern floating rate period. Journal of Finance 43, 933-948.

Murray, C. J., Papell, D. H., 2002. The purchasing power parity persistence paradigm. Journal of International Economics 56, 1-19.

Murray, C. J., Papell, D. H., 2005. Do panels help solve the purchasing power parity puzzle? Journal of Business & Economic Statistics 23, 410-415.

Papell, D. and Theodoridis, H., 1998. Increasing evidence of purchasing power parity over the current float. Journal of International Money and Finance 17, 41-50.

Park, H. J., Fuller, W. A., 1995. Alternative estimators and unit root tests for the autoregressive Process. Journal of Time Series Analysis 16, 415-429.

Pedroni, P., 2004. Panel cointegration: Asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis. Econometric Theory 20, 597-625.

Phillips, P. C. B., 1978. "Edgeworth and saddlepoint approximations in the first-order noncircular autogression. Biometrika 65, 91-98.

Rogoff, K., 1996. The purchasing power parity puzzle. Journal of Economic Literature, 34, 647-668.

Taylor, M. P., Sarno, L., 1998. The behavior of real exchange rates during the post Bretton Woods period. Journal of International Economics 46, 281-312.

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

Thomas, D. H., 1970. Some contributions to radial probability distributions, statistics, and the operational calculi. Ph.D. dissertation, Wayne State University.

Whistler, D., White, K. J., Wong, S. D., Bates, D., 2004. Shazam Econometrics Software Version 10: User's Manual. Northwest Econometrics, Vancouver.

Wu, Y., 1996. Are real exchange rates nonstationary? Evidence from panel-datatests. Journal of Money, Credit, and Banking 28, 54-63.

Footnotes

1. Our method is computationally simpler than that of Kilian and Zha (1999) and does not require the specification of a prior.

2. For general discussions of saddlepoint approximations, see Goustis and Casella (1999) and Huzurbazar (1999). Giles (2001) provides another good illustration of the accuracy of a particular saddlepoint application.

3. One or both of $\overline{\alpha}$ and $\breve{\alpha}$ may possibly take values on the boundary of the [0 , 1] interval. In this case, we can set their value(s) to 1-ε or ε ($\varepsilon \rightarrow 0$) appropriately. Then we take the limit, and the proof still holds.

4. A random vector, x, is spherically symmetric if its distribution is the same as that of Px, for all orthogonal matrices, P. The vector x is elliptically symmetric, with characteristic matix Γ, if $\Gamma^{-1/2}x$ is spherically symmetric. It is not widely recognized in the econometrics literature that many standard results require only spherical symmetry, not normality. For example, Thomas (1970) proves that the usual t and F statistics associated with the linear regression models have their usual null distributions when normality is relaxed to spherical symmetry. Similar results relating to the Durbin-Watson test statistic (and other regression statistics that are scale-invariant) are established by Kariya and Eaton (1977), and King (1979, 1980), and others. See Chmielewski (1981) for an excellent review of the associated statistical literature.

5. Some studies have found that the key determinant of the speed of PPP convergence is the nominal exchange rate (not the price), and the slow reversion of PPP is due to the slow reversion of nominal exchange rate (*e.g.*, Engel and Morley, 2001; Cheung, *et al.*, 2004). This can also explain why the empirical consensus is inconsistent with the sticky-price model.