

A SADDLEPOINT APPROXIMATION TO THE DISTRIBUTION OF THE HALF-LIFE ESTIMATOR IN A STATIONARY AUTOREGRESSIVE MODEL

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*

February 2010

Abstract

We derive saddlepoint approximations for the distribution and density functions of the half-life estimated by OLS from autoregressive time-series models. Our results are used to prove that none of the integer-order moments of these half-life estimators exist. This provides an explanation for the very large estimates of persistency, and the extremely wide confidence intervals, that have been reported by various authors – for example in the empirical economics literature relating to purchasing power parity.

Keywords: Saddlepoint approximation; half-life estimator; autoregressive model

MSC2000 classifications: 62E17; 62F25; 62M10; 62P20

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

1. Introduction

In this paper we obtain an analytic approximation to the distribution of the half-life estimator in a stationary autoregressive model, and prove that this distribution has no finite integer-order moments. The half-life, defined as the time taken for a unit shock to dissipate by 50%, is a commonly-used measure of persistence in an autoregressive time-series model. Examples of this appear in the purchasing power parity literature in economics (*e.g.*, Abuaf and Jorion, 1990; Glen 1992; Cheung and Lai, 1994, Rogoff, 1996, and others). The so-called “purchasing power parity puzzle” refers to the surprisingly large half-life estimates that have been obtained in this literature, and the fact that when confidence intervals are reported they are generally so wide as to be of little practical use. Our results provide an explanation for these empirical phenomena.

The paper is constructed as follows. Section 2 reviews some issues associated with half-life estimation. 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. Estimating the half-life of adjustment

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

$$y_t = \alpha y_{t-1} + u_t ; \quad t = 0, 1, 2, \dots, T \quad (1)$$

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:

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

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.

It is well known that the OLS estimator, $\hat{\alpha}$, is negatively biased in small samples, and the absolute bias increases with the persistence of the series. Andrews' (1993) median-unbiased estimator provides a good tool to correct the bias. Empirical studies that apply the median-

unbiased estimator (*e.g.*, Murray and Papell, 2002; Cashin and McDermott, 2003; Caporale *et al.*, 2005; Lopez *et al.*, 2004) yield estimated half-lives that are even higher than their OLS counterparts, and with confidence intervals are still so wide that no strong conclusions can be made.

Some researchers have tried to resolve these issues by replacing (1) with a nonlinear model (*e.g.*, Taylor *et al.*, 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 fruitful basis for estimating half-lives. 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.

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 in (2). 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. Based on the density function, we then prove that the moments of the half-life estimator do not exist, and we also extend these results to the general AR(p) model. This provides an explanation for the wide confidence intervals encountered in many empirical studies, and it also implies that these studies may be using 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

From (2), 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. Various studies (*e.g.*, Phillips, 1978; Lieberman, 1994) have considered the properties of $\hat{\alpha}$ in (1). In particular, Lieberman (1994) derived a saddlepoint approximation (Daniels, 1954; Goustitis and Casella, 1999; Huzurbazar, 1999) for the density and distribution functions for the OLS estimator in the AR(1) model. He expresses the OLS estimator

$\hat{\alpha}$ in (1) as:

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

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

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

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

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

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

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

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

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

where $\hat{\varepsilon} = \left(\log |\hat{A}| \right)^{\frac{1}{2}} \text{sgn}(\hat{w})$, $\hat{z} = \hat{w} \left[2 \text{tr}\{(\hat{A}^{-1} D)^2\} \right]^{\frac{1}{2}}$, $D = D(x)$; Φ and ϕ are the standard normal distribution and density functions respectively, and \hat{w} is defined by (5). The accuracy of saddlepoint approximations is well known, with recent examples being provided by Giles (2001), Chen and Giles (2008) and many others.

3.2 Density and distribution functions of the half-life estimator

A saddlepoint approximation for the density of the half-life estimator can be derived simply from that for the density of $\hat{\alpha}$, using the transformation:

$$f(\hat{h}) = f[\hat{\alpha}(\hat{h})|0 < \hat{\alpha} < 1]J = \frac{f(\hat{\alpha}(\hat{h}))J}{\text{Pr.}(0 < \hat{\alpha} < 1)}, \quad (7)$$

where $f(\hat{\alpha}(\hat{h}))$ is the density function obtained by replacing $\hat{\alpha}$ with $(0.5)^{1/\hat{h}}$ in (4); and the Jacobian is $J = (0.5)^{1/\hat{h}} \ln 2 / \hat{h}^2$.

Then $C = \text{Pr.}(0 < \hat{\alpha} < 1)$ can be calculated from (6), and the saddlepoint approximation to the density function for the half-life estimator is:

$$\hat{f}(\hat{h}) = \frac{\left\{ \text{tr}(\tilde{A}^{-1} R'_\alpha C_2 R_\alpha) \right\} |\tilde{A}|^{-1/2} (0.5)^{1/2} \ln 2}{\left[4\pi \text{tr}\{(\tilde{A}^{-1} \tilde{D})^2\} \right]^{1/2} C \hat{h}^2}, \quad (8)$$

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

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

$$\begin{aligned} \hat{F}(\hat{h}) &= P(\hat{h} < x | 0 < \hat{\alpha} < 1) \\ &= P\left(\frac{\log(0.5)}{\log(\hat{\alpha})} < x\right) / P(0 < \hat{\alpha} < 1) \\ &= P(\hat{\alpha} < (0.5)^{1/x}) / C. \end{aligned} \quad (9)$$

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

Figure 1 shows the saddlepoint approximation to the density function for \hat{h} (from (8)) for various values of α , and a sample size of 30. Figure 2 shows this density of \hat{h} for various sample sizes with $\alpha = 0.95$. We see that the density function is highly skewed to the right, and it shifts to the right and the tails become fatter as the sample size or the value of α increase. It is clear why relatively large half-life estimates have been reported frequently in the empirical literature.

[Figures 1 and 2 about here]

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

α	$T = 30$		$T = 50$	
	Point estimate (Median)	95% Confidence Interval	Point estimate (Median)	95% Confidence Interval
0.6	1.445	[0.536, 3.700]	1.409	[0.660, 2.880]
0.7	2.120	[0.746, 6.216]	2.051	[0.916, 4.529]
0.8	3.550	[1.095, 14.946]	3.379	[1.348, 8.817]
0.9	8.372	[1.900, ∞]	7.762	[2.352, 41.710]
0.97	31.786	[4.659, ∞]	30.235	[15.045, ∞]

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 the 95% confidence interval of the half-life estimator when the true data process is an AR(1) model. We see that the point estimate decreases with the sample size, and when α is no smaller than 0.9 the confidence interval is very wide. As the sample size increases, the confidence interval width decreases, but it is still quite wide.

3.3 Non-existence of moments

Further insights into the characteristics of the half-life estimator can be obtained by considering the moments of its distribution. Interestingly, 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 $u(\hat{\alpha}) = \left\{ \text{tr}(\hat{A}^{-1} R'_\alpha C_2 R_\alpha) \right\} |\hat{A}|^{-\frac{1}{2}}$ and $v(\hat{\alpha}) = \left[4\pi \text{tr}\{(\hat{A}^{-1} D)^2\} \right]^{\frac{1}{2}}$. Then

$$\begin{aligned} \hat{M}(\hat{h}) &= \int_0^1 \frac{\log(0.5) u(\hat{\alpha})}{\log(\hat{\alpha}) v(\hat{\alpha})} d\hat{\alpha} \\ &= \text{Lim}_{\varepsilon \rightarrow 0} \int_\varepsilon^{1-\varepsilon} \frac{\log(0.5) u(\hat{\alpha})}{\log(\hat{\alpha}) v(\hat{\alpha})} d\hat{\alpha}. \end{aligned}$$

As $-\infty < \hat{\alpha} < \infty$, $u(\hat{\alpha})$ 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 $\hat{\alpha} = \bar{\alpha}$, $u(\hat{\alpha})$ gets to its minimum value N and $N \neq 0$.

(ii) when $\hat{\alpha} = \tilde{\alpha}$, $v(\hat{\alpha})$ 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:

$$\begin{aligned} \hat{M}(\hat{h}) &> \text{Lim}_{\varepsilon \rightarrow 0} \int_\varepsilon^{1-\varepsilon} \frac{\log(0.5) N}{\log(\hat{\alpha}) M} d\hat{\alpha} \\ &= \log(0.5) \frac{N}{M} \text{Lim}_{\varepsilon \rightarrow 0} \int_\varepsilon^{1-\varepsilon} \frac{1}{\log(\hat{\alpha})} d\hat{\alpha} \\ &= \log(0.5) \frac{N}{M} \left[\text{Lim}_{\varepsilon \rightarrow 0} \left(\frac{\hat{\alpha}}{\log(\hat{\alpha})} \right) \Big|_\varepsilon^{1-\varepsilon} + \text{Lim}_{\varepsilon \rightarrow 0} \int_\varepsilon^{1-\varepsilon} \frac{1}{[\log(\hat{\alpha})]^2} d\hat{\alpha} \right] \\ &= \log(0.5) \frac{N}{M} \left[\infty - \text{Lim}_{\varepsilon \rightarrow 0} \int_\varepsilon^{1-\varepsilon} \frac{1}{[\log(\hat{\alpha})]^2} d\hat{\alpha} \right]. \end{aligned}$$

So, the estimated mean of the half-life estimator does not exist. ■

Based on the inversion formula, we know that

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

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.

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 none of the integer-order moments of the half-life estimator exist.

The proof follows that of Theorem 1.

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

Now consider the situation where the data follow an AR(p) process:

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

There is no explicit expression for the half-life based on the estimation of the coefficients in (10). The approach often used in practice is to re-formulate (10) into a form reminiscent of an augmented Dickey-Fuller (ADF) regression:

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

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

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

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

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

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:

$$\hat{\beta} = \frac{R' Q R}{R' G R}, \quad (13)$$

where

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

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

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

$$\hat{\beta} = \frac{v' P' Q P v}{v' P' G P v}, \quad v \sim N(0, \sigma^2 I). \quad (14)$$

As (14) and (3) are very similar we can derive the approximation to the density function of $\hat{\beta}$ as

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

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

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

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

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

where $\hat{\varepsilon} = (\log|\hat{N}|)^{\frac{1}{2}} \text{sgn}(\hat{w})$, $\hat{z} = \hat{w} \left[2\text{tr}\{(\hat{N}^{-1} L)^2\} \right]^{\frac{1}{2}}$, $L = L(x)$, Φ and ϕ are the standard normal distribution and density functions respectively, and \hat{w} is defined by (16).

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

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

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 $\text{tr}(\tilde{N}^{-1} \tilde{L}) = 0$. Using (12), and allowing for the Jacobian, the density function of the half-life

estimator in the AR(p) model is:

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

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

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

Let $K = \Pr.(-1 < \hat{\beta} < 0)$, which can be calculated from (17). Then, the saddlepoint approximation for the density function for the half-life estimator is:

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

where $\bar{L} = L(\hat{h}) = P'(Q - ((0.5)^{1/\hat{h}} - 1)G)P$, $\bar{N} = N(\hat{w}) = I - 2\hat{w}\bar{L}$ and \hat{w} satisfies $\text{tr}(\bar{N}^{-1} \bar{L}) = 0$. Based on the density function (20), we have the following results, the proofs of which again follow that of Theorem 1.

Theorem 2

Suppose that the data follow a stationary AR(p) process and satisfy the ADF equation:

$$\Delta y_t = \beta y_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta y_{t-i} + u_t, \text{ with } u_t \sim N(0, \sigma^2) \text{ and } \beta \in (-1, 1), \text{ and the half-life is defined}$$

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.

Corollary 2

Suppose that the data follow a stationary AR(p) process and satisfy the ADF equation:

$$\Delta y_t = \beta y_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta y_{t-i} + u_t, \text{ with } u_t \sim N(0, \sigma^2) \text{ and } \beta \in (-1, 1), \text{ and the half-life is defined}$$

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.

5. Robustness results

We have assumed that the data are normally distributed, but we now consider the robustness of our results to the relaxation of this assumption. King (1979; p. 121) proves, *inter alia*, that when the disturbance vector in a first-order autoregressive model follows an elliptically symmetric 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. Accordingly, we would anticipate that our own results will be robust to departures from normality, within the elliptically symmetric family of distributions. 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 Chmielewski (1981) provides an excellent review of the associated statistical literature.

Our concern here is whether the non-existence of the moments of the half-life estimator still holds under more general distributional assumptions. We can apply further results of Lieberman (1997) to establish the robustness of our earlier theorems. 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, let

$$\begin{aligned} S &= v' R'_\alpha C_1 R_\alpha v - \alpha v' R'_\alpha C_2 R_\alpha v \\ Z &= v' R'_\alpha C_2 R_\alpha v \\ B &= R'_\alpha C_2 R_\alpha . \end{aligned}$$

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

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

with the saddlepoint \hat{w} satisfying

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

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

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

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

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

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

Suppose that v has arbitrary cumulants $k^i = 0, k^{i,j}, k^{i,j,k}, \dots$, defined as follows:

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

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

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

$$\tilde{k}_2^S = \sum_{ijkl} s_{ij} s_{kl} k^{ij,kl} \quad (27)$$

$$\tilde{k}_{10} = \sum_{ij} b_{ij} k^{ij} \quad (28)$$

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

$$\tilde{k}_2^S = k_4 \sum_{ij} s_{ij}^2 + 2k_2^2 \sum_{ij} s_{ij}^2 \quad (29)$$

$$\tilde{k}_{10} = k_2 \sum_i b_{ii} \quad (30)$$

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

The approximating function in (21) is continuous on a closed interval $\hat{\alpha} \in [0, 1]$. Let

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

so

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

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). 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 Corollary 1, it is readily seen that none of the integer-order moments of \hat{h} exist in this more general case.

So, our main results hold as long as v is *i.i.d.* and the second cumulant of v is finite. When we allow the disturbances to be correlated, the situation is more complicated. However, there is still quite a large class of distributions satisfying 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. The non-existence results in Corollaries 1 and 2 are quite robust to the distributional assumption.

6. Conclusions

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 integer-order moments of such half-life estimator do not exist. These results are 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. Fundamentally, 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). Future work may be better to focus on constructing more appropriate measures of persistence, rather than just explore possible ways of improving the accuracy of the traditional half-life measure.

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: Density of the half-life estimator when $T = 30$

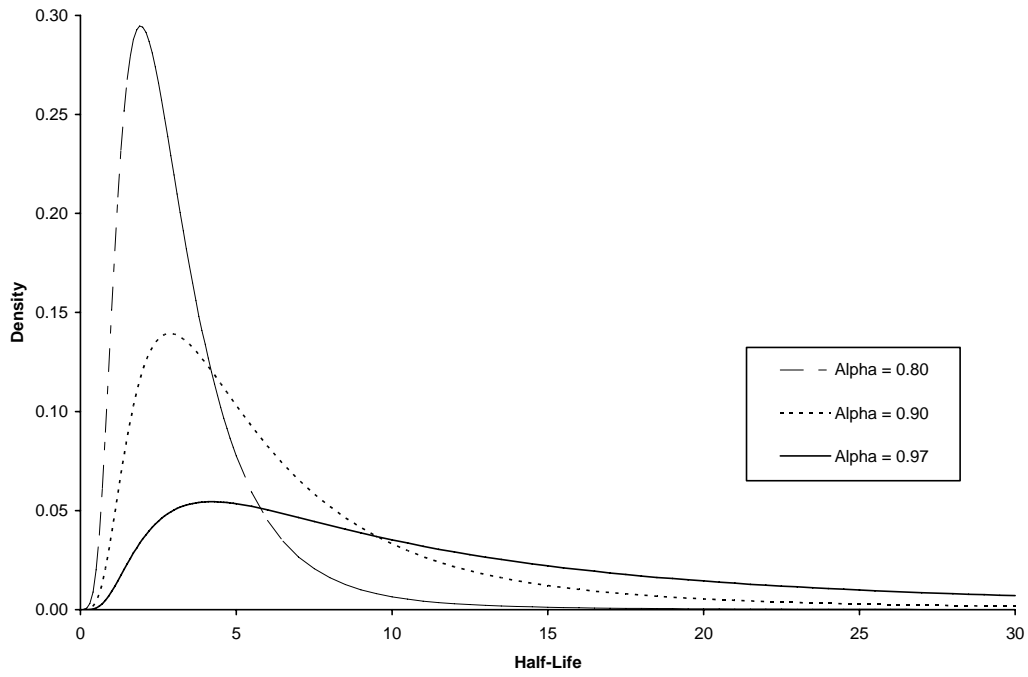
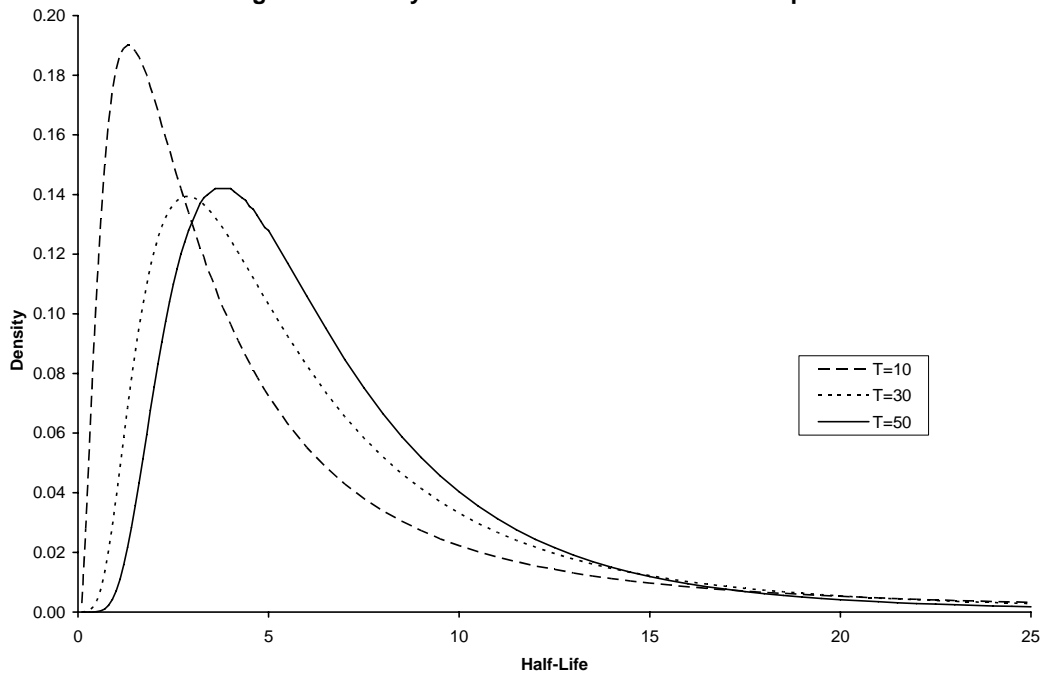


Figure 2: Density of the half-life estimator when $\alpha = 0.9$



Appendix

Here we justify assumptions (i) and (ii) used in the proof of Theorem 1. First, we prove that $N \neq 0$:

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

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

As the density exists, it follows that $\left| \hat{A} \right|^{-\frac{1}{2}} \neq 0$.

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

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

as $\left| \hat{A} \right|^{-\frac{1}{2}} = \exp\left\{ -\frac{1}{2} \sum_{t=0}^T \log(1 - 2\hat{w}d_t) \right\}$ 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:

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

and so $N \neq 0$.

Second, we prove that M is finite:

$$M = v(\tilde{\alpha}) = \left[4\pi \text{tr}\{(\hat{A}^{-1} D)^2\} \right]^{\frac{1}{2}} = \left\{ 4\pi \sum_{t=0}^T \left[d_t^2 / (1 - 2\hat{w}d_t)^2 \right] \right\}^{\frac{1}{2}},$$

so, if M is infinite, 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 is finite.

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-unbiased 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.
- Chen, Q., Giles, D. E. A. 2008. General saddlepoint approximations: Application to the Anderson-Darling test statistic. *Communications in Statistics, B* 37, 789-804.
- Cheung, Y. W., Lai, K. S., 1994. Mean reversion in real exchange rates. *Economics Letters* 46, 251-256.
- Chmielewski, M. A., 1981. Elliptically symmetric distributions: A review and bibliography. *International Statistical Review* 49, 67-74.
- 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.
- El-Gamal, M.A., Ryu, D., 2006. Short-memory and the PPP hypothesis. *Journal of Economic Dynamics and Control* 30, 361-391.
- 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.
- 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., 1994. 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.
- 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.
- Murray, C. J., Papell, D. H., 2002. The purchasing power parity persistence paradigm. *Journal of International Economics* 56, 1-19.
- 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., 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.
- Whistler, D., White, K. J., Wong, S. D., Bates, D., 2004. Shazam Econometrics Software Version 10: User's Manual. Northwest Econometrics, Vancouver.