

The Effects of Small Sample Bias in Threshold Autoregressive Models*

Yamin Ahmad †

University of Wisconsin-Whitewater

Abstract

This paper investigates Threshold Autoregressive (TAR) models that contain a limited number of observations in some regimes. Simulations show that within the context of the real exchange rate literature, parameter estimates exhibit significant small sample bias even with long time series data. These distortions create substantial power losses in attempting to identify values of coefficients from data.

JEL Classification: F47 C15 C32

Keywords: Threshold autoregressive models, small sample bias, simulation

Current Version: June 2007

* I would like to thank John Rogers for the initial premise for this paper. Any remaining errors are my own.

† Department of Economics, University of Wisconsin-Whitewater, 800 W Main St, Whitewater WI 53190, Email: ahmady@uww.edu, Homepage: <http://facstaff.uww.edu/ahmady/> Tel: (262) 472 5576, Fax: (262) 472 4683

1 Introduction

Conventional wisdom would dictate that when there are insufficient numbers of observations in data, we would obtain an imprecise and biased estimate of parameters we would wish to infer. This paper explores the impact of limited observations in estimating coefficients within a model that exhibits nonlinear dynamics, known as a threshold autoregressive (TAR) model. TAR models and its variants have become popular in several literatures.¹ Within a TAR model, the evolution of the dependent variable may follow different time series processes for different sub samples. The particular regime specific process adopted depends on the observed historical value of that variable, as does the transition from one regime to another.

Efficient estimation of coefficients in different regimes carries the implicit requirement that there are sufficient numbers of observations within each regime. If there are insufficient observations in any given regime, then the estimation methodology is subject to small sample bias and yields inefficient estimates. Using the literature on real exchange rates as an example, this paper presents some informal evidence that regimes may contain small samples and then uses simulations to examine the extent to which small sample bias affects parameter estimates within the class of threshold autoregressive models.²

2 A TAR model of the Real Exchange Rate

The standard model within the real exchange rate literature is one where the real exchange rate is assumed to follow a linear AR(1) process. Defining the log of the real exchange rate, q_t , at time period t as:

¹Two examples from different literatures include Christopoulos and León-Ledesma (2007) and Kilian and Taylor (2003). Taylor and Taylor (2004) contains an excellent review of application of TAR models to the literature on real exchange rates.

²Although the paper examines the issue within the context of the real exchange rate literature, the thesis of the paper is valid for any models which utilize threshold autoregressions or its variants.

$$q_t \equiv s_t - p_t + p_t^* \tag{1}$$

$$\text{then } q_t = \lambda q_{t-1} + u_t$$

where $u_t \sim N(0, \sigma_u^2)$, s_t , p_t and p_t^* are the logarithms of the nominal exchange rate (written as the domestic price of foreign currency), the domestic and foreign price levels respectively at time t . A TAR model of the real exchange rate utilizes frictions in order to introduce a band where persistent deviations from PPP can exist as an equilibrium feature. Various justifications for the existence of the bands include the presence of international transportation costs (Michael, Nobay & Peel, 1997); commodity points (Obstfeld & Taylor, 1997); and noise traders (Kilian and Taylor, 2003). Within the band, real exchange rates can exhibit random walk type behavior. Once outside some threshold value, price differentials are sufficiently large that arbitrage becomes profitable and so the real exchange rate is symmetrically mean reverting. Hence the entire time series process is globally stationary.

I consider the family of $TAR(p, d)$ models which are characterized by the order of autoregression, p , and an arbitrary delay parameter d . The delay parameter, d , is the particular lag order that is used to determine within which threshold regime the current observation falls. Let y_t represent the deviation from PPP, defined as:

$$s_t - p_t + p_t^* = y_t \tag{2}$$

The process y_t exhibits two types of behavior. If the historical value of the real exchange rate lies in the inner regime, i.e. $|y_{t-d}| < c$, arbitrage does not occur and the real exchange rate exhibits a random walk type behavior. In the outer regime, i.e. $|y_{t-d}| > c$, arbitrage occurs as in the standard model and we observe mean reversion towards the center of the band. I follow Obstfeld and Taylor (1997) in selecting $p = d = 1$ and write the TAR specification in terms of deviations from PPP:

$$y_t = \begin{cases} \lambda^{out} y_{t-1} + \varepsilon_t^{out} & \text{if } y_{t-1} > c; \\ \lambda^{in} y_{t-1} + \varepsilon_t^{in} & \text{if } c \geq y_{t-1} \geq -c \\ \lambda^{out} y_{t-1} + \varepsilon_t^{out} & \text{if } -c > y_{t-1} \end{cases} \quad (3)$$

where $e_t^{out} \sim N(0, \sigma_{out}^2)$, $e_t^{in} \sim N(0, \sigma_{in}^2)$. $\lambda^{in} (= 0)$ and λ^{out} represent the speeds of convergence inside and outside the thresholds respectively. From (3) note that if $|y_{t-1}| \leq c$, then λ^{in} determines the speed of adjustment, whilst λ^{out} does so if $|y_{t-1}| > c$.

3 Some Evidence of Small Sample Bias

This paper analyzes what occurs if the threshold, c is large, or if a particular data sample contains a limited number of observations within a particular regime. Here, I present some evidence that a limited number of observations may lie outside the threshold in the outer regime. Using monthly data obtained from the IMF's International Financial Statistics, I construct five real US dollar exchange rates with France and Italy from January 1973 up until December 1998 and with Canada, Japan and the UK up until November 2005. I use end of period spot dollar nominal exchange rates along with the CPIs for each country to construct the real exchange rate defined by (1). I then estimate a TAR model for the real exchange rate following the methodology outlined by Obstfeld and Taylor (1997), Tsay (1989) etc, as follows.

Having set $p = d = 1$, I estimate λ^{out} and c using grid-search methods and follow Balke and Fomby (1997) in searching over the threshold parameter, c .³ I minimize the sum of squared residuals for observations in the outer regime since I wish to provide the best possible fit for observations that lie in that regime. As an alternative I also maximize a weighted R^2 measure (weighted by the number of observations in each regime), which captures the overall goodness of fit for observations both inside and outside the thresholds. The search algorithm consists of the following steps:

³This is merely done to illustrate the main point of the paper and does not restrict generality. The results that I present are robust to higher values of p .

1. Sort $|y_t|$ and divide the interval into candidate thresholds c_k ; $k = 1, \dots, J$ ranking from the lowest to the highest value in steps of 0.001.
2. Choose c equal to the k^{th} highest value of c_k .
3. Partition the sample into observations that lie within the inner and outer regimes.
4. Perform OLS on each of the partitioned samples and calculate the residual sum of squares and weighted R^2 values.
5. Whilst $k < J$, increase k by 1 and go to step 2.
6. Locate the choice of k and c_k that achieves the objective being used. Note the associated $(\lambda^{out}, \lambda^{in}, \sigma_{out}, \sigma_{in})$.

An important point to note is that studies estimating TAR models typically restrict the set of thresholds that they search over, by starting at the 10th percentile for $|y_t|$ and ending at the 90th percentile at step 1. Thus, they eliminate partitions at step 3 which contain few observations. Since I wish to examine the distribution of λ^{out} under small samples, I include those candidate thresholds that lead to limited observations in the outer regime. Table 1 reports the results from the estimating the $TAR(1, 1)$ model.

From panel A, an unrestricted grid search over c that minimizes the residual sum of squares leads to few observations in the outer regime for all countries. The lack of observations yields inefficient estimates of λ^{out} (and σ_{out}) and renders them unreliable. With the exception of the dollar-sterling real exchange rate, none of the λ^{out} s are significant. For the inner regime, we are unable to reject the null hypothesis of a unit root in the real exchange rate for any country. Since the sum of squared residuals criterion is subject to the same distortions as the parameter estimates, additional evidence from the weighted R^2 criterion measuring the overall goodness of fit for both regimes in panel B finds the same choice of threshold values for all countries with the exception of Japan. For Japan, a smaller threshold value of 0.4 is obtained along with a unit increase in the number of observations that lie in the outer regime. However, λ^{out} is once again insignificant even though the inner regime indicates random-walk type behavior.

4 The Effect of Limited Observations

Since I seek to understand the extent to which the estimates of λ^{out} are distorted when there are a small number of observations, I conduct simulations to derive the small sample distribution for λ^{out} and the number of observations falling in the outer regime. In order to set the simulation up so that it corresponds to actual data, I use the parameters estimated for the dollar-sterling real exchange rate and set $c = 0.3$, $\sigma_{out} = \sigma_{in} = 0.025$, $\lambda^{in} = 1$. I consider three values for $\lambda^{out} \in \{0.5, 0.75, 0.9\}$ and conduct the Monte Carlo simulations as follows:

1. Select $n + b$, where n is the length of the pseudo data to be generated and b are the number of initial values to discard to avoid initial value bias.
2. Generate $i = 1, \dots, m$ trials of the data using the TAR specification in (3).
3. For each simulation, estimate the TAR(1,1) model as above using a grid search on c .
4. Construct the empirical distributions for observations falling within the outer regime and for λ^{out} .

I chose $n \in \{200, 400, 600\}$, $b = 200$ and $m = 1000$. The results are reported in table 2 and figures (1) - (4). Figure 1 depicts an arbitrarily chosen draw of the simulated data where 10 observations fall into the outer regime, and shows how it corresponds to the actual dollar-sterling real exchange rate series. The sample standard deviation of the depicted series is 0.158 compared to 0.133 for the actual dollar sterling real exchange rate. Figure 2 depicts the actual number of observations that fall outside the threshold from the simulation. Note that there are relatively few observations which lie outside the thresholds. The median number of observations falling outside the threshold is 12, with the mean equal to 13.

Table 2 reports the distribution of observations that lie outside the threshold, as well as the small sample distribution obtained for the three values of λ^{out} I examine. The limited number of observations outside the threshold leads to substantial distortions in estimates of λ^{out} . The

mean bias, measured as $E(\hat{\lambda}^{out}) - \lambda^{out}$, ranges from -0.71 to -1.23. Moreover, the bias appears to be highly significant for all estimates of λ^{out} once the dispersion of the parameter estimates are taken into account. This can be seen in figure 3, which plots the small sample distributions of λ^{out} . The magnitude of the bias is mitigated somewhat the smaller the true value of λ^{out} . However, a significant bias remains.

Figure 4 depicts the results from holding the true value of λ^{out} constant, but varying the length of the generated series which is analyzed. I examine series of length 200, 400 and 600 for $\lambda^{out} = 0.9$. As the histograms show, the number of observations at the true value of λ^{out} increases as the length of the analyzed series increases. As one would expect, time series data of greater length would have increased power in detecting the true value of λ^{out} since one would expect to obtain a larger number of instances where λ^{out} is correctly estimated and inferred. When the length of the series, $n = 200, 400$ and 600 , the true value of λ^{out} is correctly inferred approximately 1 percent, 4.5 percent and 10 percent of the time respectively. This result highlights the extensive loss in power in being able to detect the correct value of λ^{out} with a limited number of observations.

5 Conclusions

This paper explores the effect that limited observations have on coefficient estimates for regimes, within threshold autoregressive models. Using the literature on real exchange rates as an example, I highlight the proposition that some regimes may contain a limited number of observations. I utilize simulations to derive the small sample distribution for coefficient estimates and explore the effects in terms of biases and power losses that arise when estimating coefficients. I find that a small sample bias exists and is large in magnitude. This generates distortions and leads to inefficiency when estimating parameters. Moreover, it leads to substantial power loss in being able to detect and identify the true value of parameters in the data. The results of this study are intended to provide a note of caution since the findings here may have severe implications for papers that utilize TAR models, particularly where the underlying threshold value is large.

References

Balke, N. and T. Fomby (1997) “Threshold Cointegration”, *International Economic Review*, Vol. 38, No. 3, pp. 627 – 645.

Christopoulos, D. and M. León-Ledesma (2007) “A Long-Run Nonlinear Approach to the Fisher Effect”, *Journal of Money Credit and Banking*, forthcoming.

Kilian, L. and M. P. Taylor (2003) “Why is it So Difficult to Beat the Random Walk Forecast of Exchange Rates?”, *Journal of International Economics*, Vol. 60, No. 1, pp. 85-107.

Michael, P., Nobay, R. A. and Peel, D. A. (1997) “Transaction Costs And Nonlinear Adjustments In Real Exchange Rates: An Empirical Investigation”, *Journal Of Political Economy*, Vol. 105, No. 4, August 1997, pp. 862-79.

Obstfeld, M. and Taylor, A. (1997) “Nonlinear Aspects of Goods-Market Arbitrage and Adjustment: Heckscher’s commodity points revisited”, *Journal of the Japanese and International Economies*, Vol. 11, pp 441-479

Taylor, A. and M. P. Taylor (2004) “The Purchasing Power Parity Debate”, *Journal of Economic Perspectives*, Vol. 18, No. 4, pp. 135-158.

Tsay, R. (1989) “Testing and Modeling Threshold Autoregressive Processes.”, *Journal of American Statistics Association*, Vol. 84, (March 1989), pp. 231-240.

Tables

	Canada	France	Italy	Japan	UK
Panel A: Residual Sum of Squares Criterion					
Threshold Value	0.258	0.424	0.353	0.45	0.308
Outer Regime:					
# observations	4	4	4	4	7
λ^{out}	1.883 (0.249)	0.590 (0.444)	-12.751 (0.407)	0.565586 (0.837)	0.715* (0.067)
σ^{out}	0.003	0.020	0.369	0.029	0.033
Inner Regime					
# observations	389	306	306	388	385
λ^{in}	0.992 (0.452)	0.978 (0.984)	0.978 (0.739)	0.986 (0.285)	0.973 (0.105)
σ^{in}	0.012	0.026	0.025	0.029	0.025
Panel B: R² Criterion					
Threshold Value	0.258	0.424	0.353	0.4	0.308
Outer Regime:					
# observations	4	4	4	5	7
λ^{out}	1.883 (0.249)	0.590 (0.444)	-12.751 (0.407)	-0.339 (0.419)	0.715* (0.067)
σ^{out}	0.003	0.020	0.369	0.020	0.033
Inner Regime					
# observations	389	306	306	388	385
λ^{in}	0.992 (0.452)	0.978 (0.984)	0.978 (0.739)	0.986 (0.277)	0.973 (0.105)
σ^{in}	0.012	0.026	0.025	0.029	0.025

Notes:

(i) Figures in parentheses represent p-values

(ii) The p-values reported for the inner regime are the results from an (Augmented) Dickey Fuller test since we assume that the real exchange rate follows a random walk within the inner regime

(iii) * denotes significance at the 10 % level

Table 1: Summary of TAR Estimation Results

	Distribution of λ^{out}			Actual Distribution of # Observations Falling Outside c
	0.5	0.75	0.9	
sample size	400	400	400	400
# trials	1000	1000	1000	1000
Coefficient Estimate				
mean	-0.729	0.041	0.183	13
median	-0.021	-0.090	-0.087	12
Standard Deviation	5.26	5.83	6.44	9.21
Mean Bias	-1.229***	-0.709***	-0.717***	
(Significance of Bias)	(-7.385)	(-3.83)	(-3.517)	
Percentile				
min	-54.244	-60.78	-65.37	0
1	-29.829	-20.71	-14.70	0
5	-2.389	-1.24	-1.02	1
10	-0.931	-0.84	-0.76	2
25	-0.506	-0.53	-0.44	6
75	0.516	0.47	0.30	18
90	0.967	0.90	0.72	24.5
95	1.249	1.33	0.98	30
99	8.296	14.59	22.06	40
max	19.611	89.80	83.30	64

Notes:

- (i) the Standard deviation above is the standard deviation of the parameter estimates from the Monte Carlo simulation.
- (ii) *** represents significance at the 1% level

Table 2: OLS Estimates of lambda parameter in outer regime

Figures

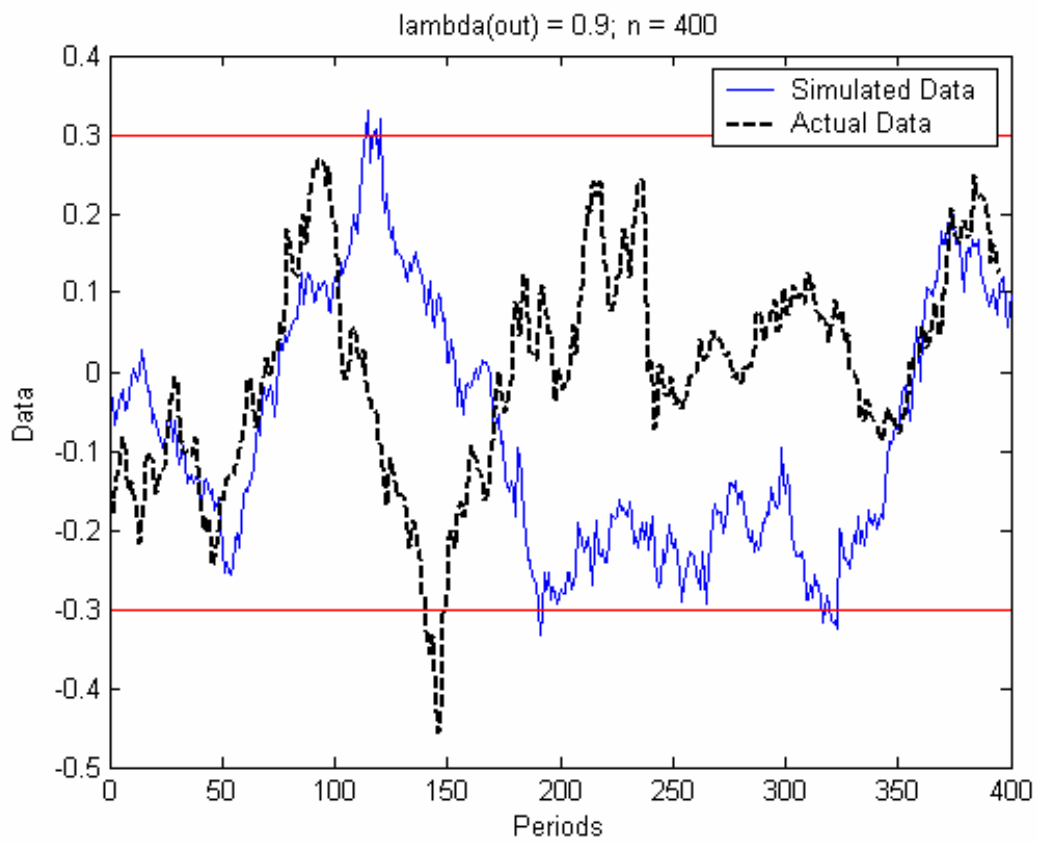


Figure 1: A Simulated Draw from a TAR(1,1) model

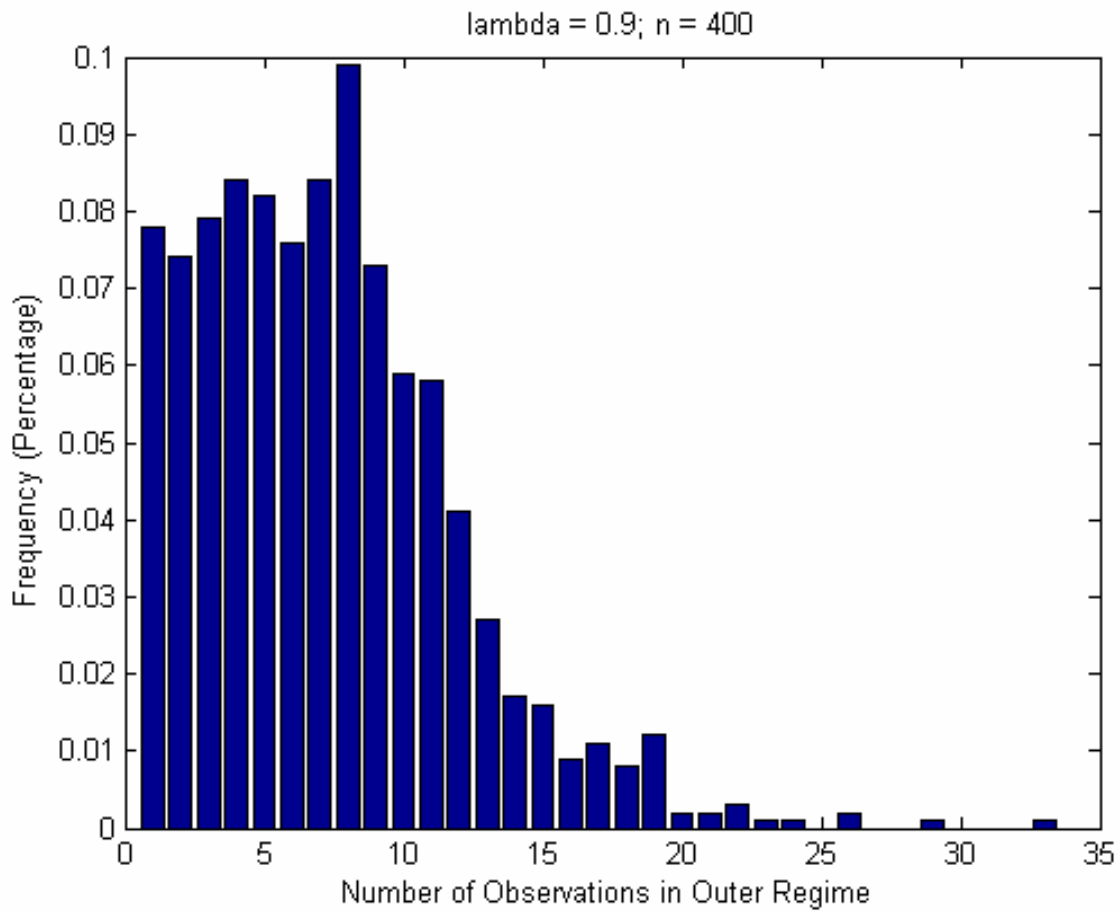


Figure 2: Distribution of the Number of Observations Falling in Outer Regime

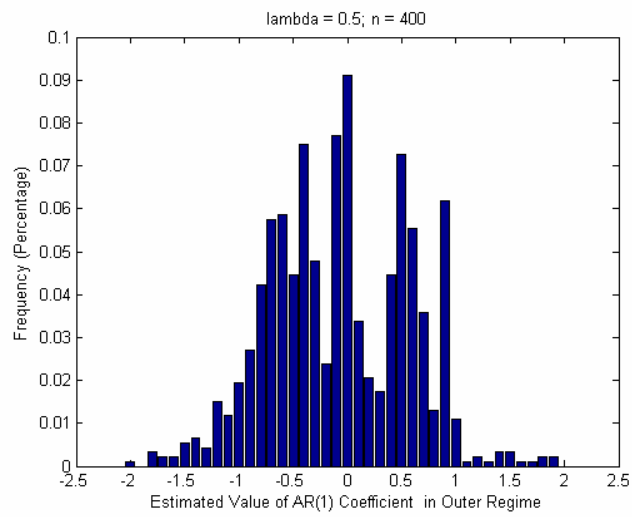


Figure 3a

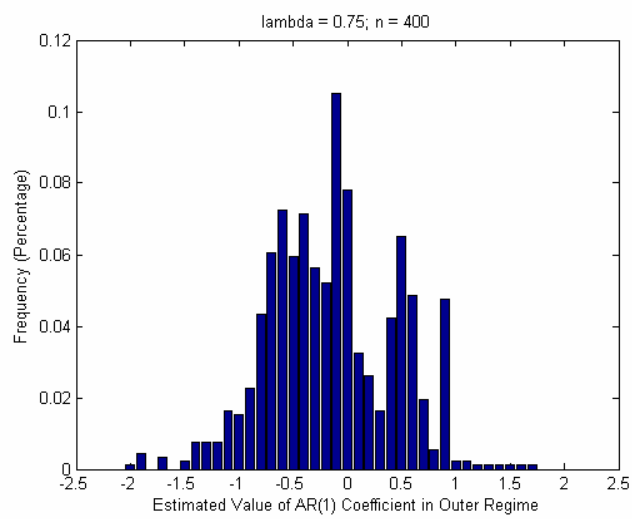


Figure 3b

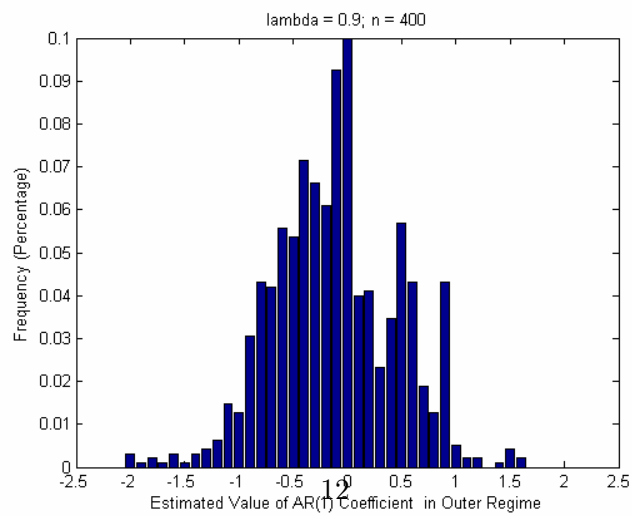


Figure 3c

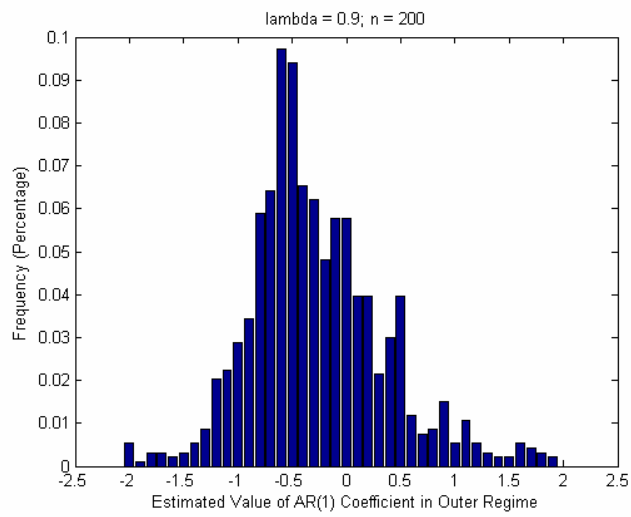


Figure 4a

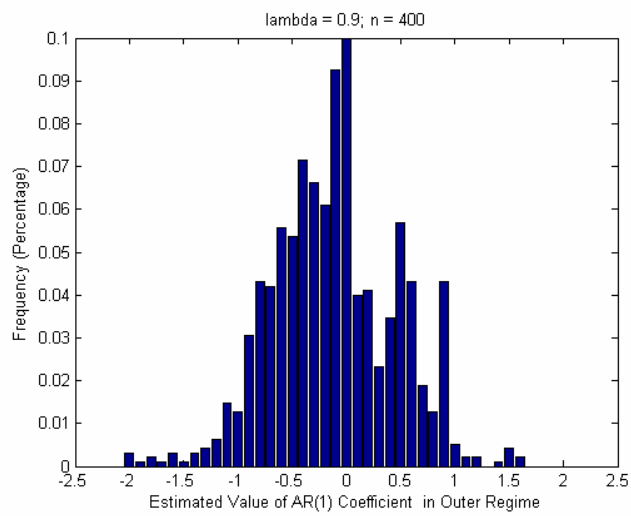


Figure 4b

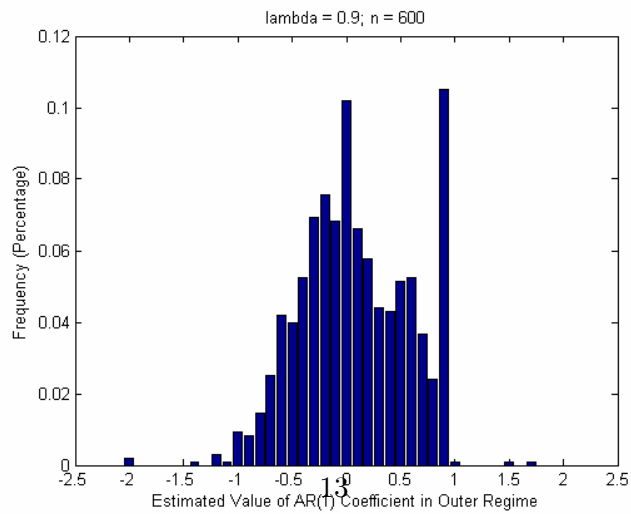


Figure 4c