Threshold GARCH Model: Theory and Application

Jing Wu*

The University of Western Ontario

Preliminary April 2010

Abstract

In this paper, we model volatility dynamics as a threshold model with an exogenous trigger, while volatility follows a GARCH process within each regime. This model can be viewed as a special case of the random coefficient GARCH model. We establish theoretical conditions, which ensure that the return process in the threshold model is strictly stationary, as well as conditions for the existence of various moments. A simulation study is further conducted to examine the finite sample properties of the maximum likelihood estimator. The simulation results reveal that the maximum likelihood estimator is approximately unbiased and consistent for modest sample sizes when the stationarity conditions hold. Furthermore, using volatility index as the threshold variable, we employ 20 stocks from Major Market Index (MMI) and find that the threshold model with an exogenous trigger fits the data well.

*I am deeply indebted to my supervisor, Professor John Knight, for his excellent guidance and continuous support and grateful to Professor Martijn van Hasselt and Youngki Shin for their advice. I also thank seminar participants at UWO and my colleagues for valuable comments. I am responsible for all errors.
1 Introduction

Modeling the temporal dependencies in the conditional variance of financial time series has been the interest of many economists and financial analysts. The most popular approaches are the ARCH model introduced by Engle (1982) and its extension GARCH model by Bollerslev (1986). To capture the striking feature that asset prices move more rapidly during some periods than others, a regime switching framework has been brought into ARCH and GARCH models. A widely used class of regime switching models is the hidden Markov model, which assumes that states of the world are unknown. While estimation is not difficult, these models often fail to generate accurate predictions due to the unknown state in the future. In this paper we employ a different type of regime switching models – the threshold model – to describe the conditional variance process. In this threshold model, the state of the world which is determined by an observable threshold variable and therefore known, while conditional variance follows a GARCH process within each state. This model can be viewed as a special case of the random coefficient GARCH models. First, we examine the theoretical and empirical properties of the threshold model with an exogenous threshold variable. We establish theoretical conditions, which ensure that the return process in the threshold model is strictly stationary, as well as conditions for the existence of various moments. A simulation study is further conducted to examine the finite sample properties of the maximum likelihood estimator. The simulation results reveal that the maximum likelihood estimator is approximately unbiased and consistent for modest sample sizes\(^1\) when the stationarity conditions hold. Furthermore, using volatility index as the threshold variable, we employ 20 stocks from Major Market Index (MMI) and find that the threshold model with an exogenous trigger fits the data well. We also explore the properties of the threshold GARCH model when the threshold variable is endogenous through simulation studies.

1.1 Regime Shifts in Conditional Variance

In most widely used GARCH models the conditional variance is defined as a linear function of lagged conditional variances and squared past returns. Formally, let \( r_t \) be a sequence of returns, \( \varepsilon_t \) be a series of innovations that are usually assumed to be independent identically distributed (i.i.d.)

\(^1\)We have estimated return data according to sample sizes ranging from 500, 1000, to 2000.
zero-mean random variable, $\sigma_t^2$ be the variance of $r_t$ given information at time $t$, the GARCH(p,q) model for returns $r_t$ is defined as follows:

$$r_t = \sigma_t \varepsilon_t \quad \sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 + \sum_{j=1}^{q} \alpha_j r_{t-j}^2$$

where $p, q = 0, 1...$ are integers, $\alpha_0 > 0, \alpha_j \geq 0, \beta_i \geq 0, i = 1, ...p, j = 1, ...q$, are model parameters.

Though these models have been proved to be adequate for explaining the dependence structure in conditional variances, they contain several important limitations, one of which is that they fail to capture the stylized fact that conditional variance tends to be higher after a decrease in return than after an equal increase. In order to account for this asymmetry many alternative models have been proposed. The exponential GARCH (EGARCH) introduced by Nelson (1991) specifies the conditional variance in logarithmic form$^2$:

$$\ln \sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \beta_i \ln \sigma_{t-i}^2 + \sum_{k=1}^{q} \alpha_k \left[ \theta Z_{t-k} + \gamma (|Z_{t-k}| - \frac{2}{\pi}) \right] \quad Z_t = \mu_t / \sigma_t$$

The threshold GARCH (TGARCH) model proposed by Zakoian (1991) and similar GJR GARCH model studied by Glosten, Jagannathan, and Runkle (1993) define the conditional variance as a linear piecewise function. In TGARCH(1,1),

$$\sigma_t^2 = \omega + \alpha r_{t-1}^2 + \delta D_{t-1} r_{t-1}^2 + \beta \sigma_{t-1}^2$$

$$D_{t-1} = \begin{cases} 1 & r_{t-1} < 0 \\ 0 & r_{t-1} \geq 0 \end{cases}$$

$^2$The model takes the asymmetry into account while keeping the linear function form of conditional variance.
More details of such alternative models can be found in the survey of GARCH models by Bollerslev, Chou, and Kroner (1992). The above alternative models are able to capture some stylized facts better than the GARCH model. However there is no evidence that any alternative model consistently outperforms the GARCH model, for example Hansen and Lunde (2005) claim that nothing beats a GARCH (1,1) in the analysis of the exchange rate data.

The TGARCH and GJR-GARCH models also relax the linear restriction on the conditional variance dynamics. Questioning the common finding of a high degree of persistence to the conditional variance in GARCH model, Lamoureux and Lastrapes (1990) suggest that such high persistence level may be spurious if there are regime shifts in the volatility process. From then both ARCH and GARCH models have been implemented with regime switching (RS) framework. The early RS applications, such as Hamilton and Susmel (1994), only allow a Markov-switching ARCH model to describe the conditional variances. Gray (1996) and Klaassen (2002), on the other hand, develop a generalized Markov-switching model, in which GARCH process in conditional variance is permitted in each regime. Comparing to the popular Markov-switching models, threshold models have clear conceptual advantages while receive less attention. Knight and Satchell (2009) derive theoretical conditions for the existence of stationary distributions for the threshold models. Based on their work it is now convenient to apply the threshold models as regime-switching models. To account for the possible structural changes in the conditional variances, this paper uses a threshold model with exogenous triggers to describe regime switches in the conditional variance process. We simply assume 2 regimes for the conditional variance, which follows a GARCH process within each regime. Different from Markov-switching models, the regimes are observable because their shifts are triggered by exogenous variables. We just need to estimate the threshold value. The model is more complex since the parameters governing the conditional variances are changing over time, however still flexible in the sense that single regime is a possible outcome in estimation procedure.

1.2 Exogenous and endogenous threshold variable

In addition to incorporating the nonlinearity in the threshold GARCH model, the threshold or trigger variable takes into account the effect of correlation between conditional variance and other

\[3\]If the estimated threshold value is the minimum or maximum of the exogenous trigger.
observed variables that represent trading activities. The use of the threshold model is particularly motivated by the volatility volume relationship.

At the time of advancing the volatility modeling, an extensive study on stock return volatility-volume relation has been developed. As mentioned in Poon and Granger (2003) the volatility-volume research may lead to a new and better way for modeling returns distribution. The early works on the relationship between stock returns and trading volume that summarized in Karpoff (1987) show that volume is positively related to the absolute value of the price change. Later works further identify the positive contemporaneous correlation between return volatility and volume (Gallant, Rossi, and Tauchen (1992), and Lamoureux and Lastrapes (1990)).

The empirical works establish various relationships between stock returns and trading volume, yet there is no consensus on how to model the underlying generating process theoretically. The favored theoretical explanation of positive price-volume correlation is the mixture of distribution hypothesis (MDH), which states that the stock returns and trading volume are driven by the same underlying latent information variable (Clark (1973), Epps and Epps (1976), Tauchen and Pitts (1983), Andersen (1996), and Bollerslev and Jubinski (1999)). One encouraging attempt is Andersen’s (1996) MDH model in which the joint dynamics of returns and volume are generalized and estimated with a result of significant reduction in the volatility persistence.

More interestingly, recent findings suggest that the size of the trading volume, more specifically the above average volume has significant effect on conditional variance (Wagner and Marsh 2004). Intuitively, the price changes in a stock market can be regarded as a response to arrivals of information, while the volume of shares traded reflects the arrival rate of information. As mentioned in many research stock prices experience volatile period with high intensity of information arrivals and tranquil period accompanied by moderate trading activities. If we assume that the volatility follows different processes in different regimes, obviously volume provides information about which regime the volatility is in.

The established volatility-volume relation motivates the use of volume as the trigger variable in our threshold GARCH model. Since volume and volatility are highly correlated, volume must be treated as an endogenous threshold variable. Nevertheless, other variables that reflect trading activities can also be accommodated. In this paper we choose the Chicago Board Options Exchange (CBOE) Volatility Index (VIX) as an exogenous threshold variable since it is a measure of market expectations of near-term volatility and therefore has almost no correlation with current
volatility but provides information on the state of the current volatility. The VIX is calculated and disseminated in real-time by CBOE since 1993. It is a weighted blend of prices for a range of options on the S&P 500 index. The formula uses a kernel-smoothed estimator that takes as inputs the current market prices for all out-of-the-money calls and puts for the front month and second month expirations. The goal is to estimate the implied volatility of the S&P 500 index over the next 30 days.

Our paper is organized as follows. Section 2 introduces the model and derives the stationarity conditions. Section 3 provides the simulation study. Section 4 discusses the empirical findings, and section 5 concludes.

2 Model

2.1 Introduction

The threshold GARCH model we study in this paper is defined as follows:

\[ r_t = \sigma_t \varepsilon_t \]
\[ \sigma_t^2 = \omega_{st-1} + \alpha_{st-1} r_{t-1}^2 + \beta_{st-1} \sigma_{t-1}^2 \]

where \( r_t \) is the series of demeaned returns and \( \sigma_t^2 \) is the conditional variance of returns given time \( t \) information. We assume that the sequence of innovations \( \varepsilon_t \) follow the independent and identical distribution with mean 0 and variance 1: \( \varepsilon_t \sim iidD(0, 1) \). The parameters \( \{\omega_{st}, \alpha_{st}, \beta_{st}\} \) in the conditional variance equation depend on a threshold variable \( y_t \):

\[ \sigma_t^2 = \omega_0 + \alpha_0 r_{t-1}^2 + \beta_0 \sigma_{t-1}^2 \quad \text{if} \ y_{t-1} \leq y_0 \]
\[ \sigma_t^2 = \omega_1 + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \quad \text{if} \ y_{t-1} > y_0 \]

where the state or regime of the world \( S_t \) is determined by threshold variable \( y_{t-1} \) which can be treated as exogenous or endogenous and threshold value \( y_0 \) with \( E(S_t) = E(y_{t-1} > y_0) = \pi \). To simplify the theoretical derivation, we assume the threshold variable is independent of \( \sigma_t^2 \).
As in the standard GARCH(1,1) model we impose the non-negative constrains on all parameters to ensure the conditional variance to be non-negative. However, the conventional stationary conditions for GARCH model may not apply here. Since the conditional variance can fall into 2 different regimes, it is possible that conditional variance is not stationary in one regime but strictly stationary in another regime.

The conditional variance dynamics in the threshold GARCH model we define above is similar to a threshold AR (TAR) model. Knight and Satchell (2009) derive the stationarity conditions for TAR model following the work of Quinn (1982). We follow Knight and Satchell (2009) in deriving the stationarity conditions for the conditional variance and the return series accordingly. Proposition 1 gives conditions for the existence of stationary solution of return process as well as the existence of the mean in the threshold GARCH model. We also examine the conditions for the existence of higher order moments. Proposition 2 provides conditions for the return process to have a stationary variance and Proposition 3 presents conditions for the existence of the fourth moment. Since the return processes experience low autocorrelation but squared returns are highly correlated, we are also interested in examining the theoretical autocorrelation structure of the squared return. Proposition 4 expresses the formulas for the squared return autocovariance and autocorrelation functions.

2.2 Stationary Return Process
2.2.1 Mean and Variance Stationarity Conditions

Given the assumptions that $\varepsilon_t$ is iid distributed variable with $D(0, 1)$ and is independent of $\sigma_t$, it’s easy to see that the return series is mean stationary with $E(r_t) = 0$. To simplify the expression of higher order moments we further assume that $\varepsilon_t \sim iid N(0, 1)$. Thus the unconditional variance and the fourth moment of return are given by $E(r_t^2) = E(\sigma_t^2)$ and $E(r_t^4) = 3E(\sigma_t^4)$. Obviously to examine the stationarity of return we need to check the first and second moments of the conditional variance $\sigma_t^2$. The following propositions give the conditions under which the stationary distribution of return, the stationary variance, the finite fourth moment of return process, and stationary covariance exist. Proofs of the propositions are provided in Appendix A.
PROPOSITION 1. The return series has a stationary solution if \( \omega_0 < \infty, \omega_1 < \infty, \) and:

\[
[(\alpha_0 + \beta_0)(1 - \pi) + (\alpha_1 + \beta_1)\pi] < 1
\]

The stationary mean is given by:

\[
E(r_t) = E\left(\sqrt{c_0 + c_1 B_{t-1} + \sum_{n=1}^{\infty} (c_0 + c_1 B_{t-1-n}) S_n(t) \varepsilon_t}\right) = 0
\]

\[\blacksquare\]

We can now examine the first order stationarity conditions for the conditional variance process in our threshold GARCH model.

PROPOSITION 2. The return series will be variance stationary if \( \omega_0 < \infty, \omega_1 < \infty, \) and:

\[
[(\alpha_0 + \beta_0)(1 - \pi) + (\alpha_1 + \beta_1)\pi] < 1
\]

Then the stationary variance is given by:

\[
Var(r_t) = \sigma^2 = \frac{\omega_0(1 - \pi) + \omega_1 \pi}{1 - [(\alpha_0 + \beta_0)(1 - \pi) + (\alpha_1 + \beta_1)\pi]}
\]

\[\blacksquare\]
2.2.2 Higher Order Moments and Covariance Stationary Conditions

Examine the second moment of \( \sigma_t^2 \), we will obtain the fourth moment of returns.

**PROPOSITION 3.** If and only if the following conditions hold:

\[
\omega_0 < \infty, \omega_1 < \infty, [(\alpha_0 + \beta_0)(1 - \pi) + (\alpha_1 + \beta_1)\pi] < 1 \text{ and } A = [(2\alpha_0^2 + (\alpha_0 + \beta_0)^2)(1 - \pi) + (2\alpha_1^2 + (\alpha_1 + \beta_1)^2)\pi] < 1
\]

The fourth moment of the stationary distribution exists for the return process in our threshold model and is given by

\[
E(r_t^4) = 3E(\sigma_t^4) = 3\gamma_2 \frac{1 + a_0 + b_0}{(1 - A)(1 - (a_0 + b_0))} + 3\gamma_1 (1 - (a_0 + b_0)) + 2c_0(a_1 + b_1)(1 - A)
\]

Using the results from the first and second moments of \( \sigma_t^2 \), we can now derive the formulas for the autocovariance and autocorrelation functions of squared residuals:

\[
\gamma(k) = E(r_t^2 - \sigma^2)(r_{t-k}^2 - \sigma^2) \text{ and } \rho(k) = \frac{\gamma(k)}{\gamma(0)}
\]

**PROPOSITION 4.** If the conditions in proposition 3 hold, and let \( \gamma(k) = \text{Cov}(r_t^2, r_{t-k}^2) \) and \( \rho(k) = \frac{\text{Cov}(r_t^2, r_{t-k}^2)}{\text{Var}(r_t^2)} \). Then, for all \( k \geq 2 \),

\[
\gamma(k) = (a_0 + b_0)\gamma(k - 1) \text{ and for all } k \geq 1 \rho(k) = (a_0 + b_0)^{k-1}\rho(1)
\]

where

\[
\rho(1) = \frac{c_0^2[2a_0 - 2a_0b_0a_0 + A_0(2a_1^2 + (a_1 + b_1)^2)\pi(1 - \pi)]}{c_0^2(2 + A - 3A_0^2 + 3c_1\pi(1 - \pi)[c_1(1 - A_0) + 2c_0(a_1 + b_1)(1 - A)]}
\]

\[
\quad + \frac{(3a_0 + b_0(1 + 2A_0))\pi(1 - \pi)[c_1^2 + 2c_0c_1(a_1 + b_1)(1 - A)]}{c_0^2(2 + A - 3A_0^2 + 3c_1\pi(1 - \pi)[c_1(1 - A_0) + 2c_0(a_1 + b_1)(1 - A)]}
\]
with

\[ A = (2\alpha_0^2 + (\alpha_0 + \beta_0)^2)(1 - \pi) + (2\alpha_1^2 + (\alpha_1 + \beta_1)^2)\pi \]

\[ A_0 = a_0 + b_0 = (\alpha_0 + \beta_0)(1 - \pi) + (\alpha_1 + \beta_1)\pi \]

3 A Simulation Study

3.1 The Range of Parameters under Stationary Conditions

We refer \( \pi \) as the probability of volatility process in regime 2, \( \alpha_0, \beta_0 \) as parameters in regime 1 and \( \alpha_1, \beta_1 \) as parameters in regime 2 respectively. From the stationary conditions derived in previous section, we note that since the parameter \( \pi \) enters into the conditions, the sum of the parameter values in each regime is no longer required to be less than one. For example to have a stationary distribution of return, we just need a weighted sum of the sums of parameters in two regimes to be less than one:

\[ [(\alpha_0 + \beta_0)(1 - \pi) + (\alpha_1 + \beta_1)\pi] < 1. \]

To examine the effect of \( \pi \) on the range of stationary areas, we graph the stationary areas of parameters in one regime based on different \( \pi \) values for the fixed parameter values in another regime. We discuss the stationary areas of \( \alpha_1 \) and \( \beta_1 \) when \( \pi \) varies from 0.1, 0.5, to 0.9 for four sets of parameter values of \( \alpha_0 \) and \( \beta_0 \).\(^4\)

According to the stationary conditions we derived in last section, the return series will have a stationary distribution if \((\alpha_0 + \beta_0)(1 - \pi) + (\alpha_1 + \beta_1)\pi < 1\), and the fourth moment of return series exists if \((2\alpha_0^2 + (\alpha_0 + \beta_0)^2)(1 - \pi) + (2\alpha_1^2 + (\alpha_1 + \beta_1)^2)\pi < 1\). It’s easy to verify that if \( \pi \) increases, the weight of \((\alpha_1 + \beta_1)\) increases, therefore the range of \((\alpha_1 + \beta_1)\) will decrease. In Figure 1 the areas below the straight lines satisfy the stationary distribution restriction, while the areas below the curves fulfill the requirement for existence of finite fourth moment. For example in the first graph from Figure 1, the sum of parameters in regime one is one, therefore the restriction for stationary distribution requires the sum of parameters in regime

\(^4\)We set \( \{\alpha_0, \beta_0\} = \{0.25, 0.75\}, \{0.25, 0.5\}, \{0.25, 0.25\}, \{0.25, 0\} \) respectively.
two to be less than 1 regardless of the value of \( \pi \), therefore only one line shows up on the graph. On the other hand, the restrictions for existence of fourth moment are described by two curves since given \((\alpha_0 + \beta_0) = 1\), there is no values of \( \alpha_1 \) and \( \beta_1 \) that satisfy such conditions when \( \pi = 0.1 \).

Figure 1: The Stationary Areas of \( \alpha_1 \) and \( \beta_1 \) given \( \{\alpha_0, \beta_0\} = \{0.25, 0.75\}, \{0.25, 0.5\}, \{0.25, 0.25\}, \{0.25, 0\}\)

### 3.2 The Simulated Paths of Return Series

In the previous section we derive the stationarity conditions for the return series described by our threshold GARCH model. Now we proceed a simulation study to examine the estimation performance of this model under different stationary conditions.
For the simulation study, we choose 3 sets of parameters for $\pi = 0.1, 0.5, 0.9$ respectively. The value of $\omega_0$ and $\omega_1$ are set to be 0.02 and 0.01 for all cases. The values of $\alpha_0$ and $\beta_0$ are fixed at 0.25 and 0.5, then $\beta_1$ is selected from the different regions in the stationary areas given $\alpha_1 = 0.25$ as shown in the second graph from Figure 1. We choose the parameters in such way that the regime 1 is always stationary, based on different probabilities with which conditional variance shifts to regime 2, we could have a non-stationary regime 2 but the whole process is still stationary. For the non-stationary case, we choose the parameter values just on the boundary to be non-stationary.

Case 1, $\pi = 0.1$:

- Stationary with 4th moment $\{\alpha_1, \beta_1\} = \{0.25, 1.5\}$
- Stationary $\{\alpha_1, \beta_1\} = \{0.25, 2.5\}$
- Non-stationary $\{\alpha_1, \beta_1\} = \{0.25, 3\}$

Case 2, $\pi = 0.5$:

- Stationary with 4th moment $\{\alpha_1, \beta_1\} = \{0.25, 0.75\}$
- Stationary $\{\alpha_1, \beta_1\} = \{0.25, 0.9\}$
- Non-stationary $\{\alpha_1, \beta_1\} = \{0.25, 1\}$

Case 3, $\pi = 0.9$:

- Stationary with 4th moment $\{\alpha_1, \beta_1\} = \{0.25, 0.7\}$
- Stationary $\{\alpha_1, \beta_1\} = \{0.25, 0.75\}$
- Non-stationary $\{\alpha_1, \beta_1\} = \{0.25, 0.9\}$

The data generating process for the stationary case in first sets of parameters is described as follows:

\[ r_t = \sigma_t \varepsilon_t \]

\[ \sigma_t^2 = \begin{cases} 
0.02 + 0.25r_{t-1}^2 + 0.5\sigma_{t-1}^2 & \text{if } y_{t-1} \leq y_0 \\
0.01 + 0.25r_{t-1}^2 + 1.5\sigma_{t-1}^2 & \text{if } y_{t-1} > y_0 
\end{cases} \]

$\varepsilon_t$, $y_t$ is drawn independently from standard normal distribution, $y_0$ is chosen in a way such that $E(S_t) = E(y_t > y_0) = 0.1$, and $\sigma_0$ is set to 0. We generate 5000 observations using the specified parameters, to eliminate the possible initial value effect, we drop first 2000 observations.

The paths of return series depend crucially on the parameters in volatility process. The following figure shows the stationary paths of return series given that parameters are specified in Case
3.3 The Performance of MLE Estimator

In this section we examine the performance of maximum likelihood estimator. Given that the return series is conditionally normally distributed, the log likelihood function for a sample of T observations is:

\[ \ln L_T(\theta) = -\frac{1}{2} \sum_{t=1}^{T} \ln \sigma_t^2 - \frac{1}{2} \sum_{t=1}^{T} \frac{r_t^2}{\sigma_t^2} \]

where \( \theta = \{\omega_0, \omega_1, \alpha_0, \alpha_1, \beta_0, \beta_1\} \). We know that to estimate \( \theta \), we need to estimate the threshold value \( y_0 \) so that the above likelihood function can be formulated. Here we estimate \( y_0 \) by grid search, the threshold variable \( y_t \) is sorted and for each possible threshold value \( y_0 \) we calculate the corresponding likelihood and the estimated threshold value is the one which maximizes the likelihood:

\[ \]
\[ \hat{\theta}(y_0) = \text{argmax} T^{-1} \ln L \]

The asymptotic theory for the maximum likelihood estimates of the parameters of the threshold GARCH model gives rise to difficulties because of the non-differentiability due to the threshold. Therefore we conduct a simulation study to analyze the finite sample properties of the maximum likelihood estimator.

Firstly, we present the estimation results for 3 sets of parameters in Case 1 when \( \pi = 0.1 \), considering the sample sizes for 500, 1000, and 2000. The MSE is defined as mean squared errors of estimates from true parameter values \( MSE = \frac{1}{T} \sum (\hat{\theta} - \theta)^2 \) for \( \theta = \{ \omega_0, \omega_1, \alpha_0, \alpha_1, \beta_0, \beta_1 \} \). The results are based on 1000 replications. For simplicity we estimate the threshold value by searching over the 19 grid points range from 5% percentile to 95% percentile points of threshold variable.

Table 1: The MLE estimates for parameters in Case 1

<table>
<thead>
<tr>
<th>Par</th>
<th>True Value</th>
<th>Estimate</th>
<th>MSE</th>
<th>True Value</th>
<th>Estimate</th>
<th>MSE</th>
<th>True Value</th>
<th>Estimate</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \omega_0 )</td>
<td>0.02</td>
<td>.0210</td>
<td>.0001</td>
<td>0.02</td>
<td>.0209</td>
<td>.0000</td>
<td>0.02</td>
<td>.0205</td>
<td>.0001</td>
</tr>
<tr>
<td>( \omega_1 )</td>
<td>0.01</td>
<td>.0072</td>
<td>.0045</td>
<td>0.01</td>
<td>.0096</td>
<td>.0077</td>
<td>0.01</td>
<td>.0114</td>
<td>.0094</td>
</tr>
<tr>
<td>( \alpha_0 )</td>
<td>0.25</td>
<td>.2389</td>
<td>.0076</td>
<td>0.25</td>
<td>.2487</td>
<td>.0051</td>
<td>0.25</td>
<td>.2508</td>
<td>.0077</td>
</tr>
<tr>
<td>( \alpha_1 )</td>
<td>0.25</td>
<td>.2711</td>
<td>.1194</td>
<td>0.25</td>
<td>.3077</td>
<td>.2158</td>
<td>0.25</td>
<td>.3356</td>
<td>.2801</td>
</tr>
<tr>
<td>( \beta_0 )</td>
<td>0.5</td>
<td>.4905</td>
<td>.0232</td>
<td>0.5</td>
<td>.4908</td>
<td>.0077</td>
<td>0.5</td>
<td>.4725</td>
<td>.0136</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>1.5</td>
<td>1.4712</td>
<td>.8277</td>
<td>2.5</td>
<td>2.3337</td>
<td>1.0359</td>
<td>3</td>
<td>2.6001</td>
<td>1.5691</td>
</tr>
<tr>
<td>( \omega_0 )</td>
<td>0.02</td>
<td>.2052</td>
<td>.0000</td>
<td>0.02</td>
<td>.2004</td>
<td>.0000</td>
<td>0.02</td>
<td>.2038</td>
<td>.0131</td>
</tr>
<tr>
<td>( \omega_1 )</td>
<td>0.01</td>
<td>.0097</td>
<td>.0015</td>
<td>0.01</td>
<td>.0107</td>
<td>.0025</td>
<td>0.01</td>
<td>.0136</td>
<td>.0032</td>
</tr>
<tr>
<td>( \alpha_0 )</td>
<td>0.25</td>
<td>.2474</td>
<td>.0021</td>
<td>0.25</td>
<td>.2527</td>
<td>.0023</td>
<td>0.25</td>
<td>.2555</td>
<td>.0045</td>
</tr>
<tr>
<td>( \alpha_1 )</td>
<td>0.25</td>
<td>.2569</td>
<td>.0464</td>
<td>0.25</td>
<td>.2843</td>
<td>.0880</td>
<td>0.25</td>
<td>.3160</td>
<td>.1246</td>
</tr>
<tr>
<td>( \beta_0 )</td>
<td>0.5</td>
<td>.4949</td>
<td>.0042</td>
<td>0.5</td>
<td>.4929</td>
<td>.0037</td>
<td>0.5</td>
<td>.4776</td>
<td>.0093</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>1.5</td>
<td>1.4762</td>
<td>.2526</td>
<td>2.5</td>
<td>2.3752</td>
<td>.4346</td>
<td>3</td>
<td>2.6380</td>
<td>.9153</td>
</tr>
<tr>
<td>( \omega_0 )</td>
<td>0.02</td>
<td>.2020</td>
<td>.0000</td>
<td>0.02</td>
<td>.2000</td>
<td>.0000</td>
<td>0.02</td>
<td>.2016</td>
<td>.0000</td>
</tr>
<tr>
<td>( \omega_1 )</td>
<td>0.01</td>
<td>.0094</td>
<td>.0007</td>
<td>0.01</td>
<td>.0100</td>
<td>.0012</td>
<td>0.01</td>
<td>.0127</td>
<td>.0018</td>
</tr>
<tr>
<td>( \alpha_0 )</td>
<td>0.25</td>
<td>.2479</td>
<td>.0011</td>
<td>0.25</td>
<td>.2513</td>
<td>.0016</td>
<td>0.25</td>
<td>.2525</td>
<td>.0038</td>
</tr>
<tr>
<td>( \alpha_1 )</td>
<td>0.25</td>
<td>.2519</td>
<td>.0185</td>
<td>0.25</td>
<td>.2700</td>
<td>.0376</td>
<td>0.25</td>
<td>.2910</td>
<td>.0568</td>
</tr>
<tr>
<td>( \beta_0 )</td>
<td>0.5</td>
<td>.4993</td>
<td>.0022</td>
<td>0.5</td>
<td>.4952</td>
<td>.0031</td>
<td>0.5</td>
<td>.4780</td>
<td>.0096</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>1.5</td>
<td>1.4917</td>
<td>.1096</td>
<td>2.5</td>
<td>2.4091</td>
<td>.2396</td>
<td>3</td>
<td>2.6786</td>
<td>.6989</td>
</tr>
</tbody>
</table>

Table 1 presents the estimation results for 3 sets of parameters in Case 1. When \( \pi = 0.1 \), since
the probability that the conditional variance changes to regime 2 is small, we just need the sum of parameters to be less than 3.25 to fulfill the requirement for stationary distribution. In all 3 sets of parameters, whether they are strict stationary with finite fourth moment, strict stationary, or non-stationary, the MLE estimator appears to converge to the true parameter value since the mean values of the estimates approach to the true parameter values when the sample size increases from 500 to 2000. The MSE decreases when sample size increases. We notice that the MSE for $\alpha_1$ and $\beta_1$ are substantially larger than that of $\alpha_0$ and $\beta_0$, this may be caused by the nature of non-stationarity in the corresponding regime and small probability to enter that regime.

Figure 3-5 provide the estimated density of MLE estimates summarized in the above table. The estimated density is computed using kernel smoothing method. The MLE estimates are approximately unbiased and consistent even the stationarity condition is violated. As sample size increases from 500 to 2000, the MLE estimates become more efficient with smaller variances and more concentrated around true parameters. We present the density estimates for sample size of 500, 1000, and 2000 in dotted line, dashed line, and dotted and dashed line respectively, while the true parameter values are given by the solid line.
Figure 3: Kernel Smoothing Density Estimates of MLE for Stationary Returns with Fourth Moment when $\pi = 0.1$
Figure 4: Kernel Smoothing Density Estimates of MLE for Stationary Returns without Fourth Moment when $\pi = 0.1$.
Similar results are obtained for other 2 cases and reported in Appendix B. Table 5 presents the estimation results for 3 sets of parameters in Case 2. MLE estimators are still consistent and efficient as sample size increases. We also observe that the MSE of estimates in each regime are not substantially different as reported in Table 1. It may be caused by the fact that probabilities of conditional variance in each regime are equal, and we also expect higher MSE for estimates in non-stationary regime. Table 6 presents estimation results in Case 3. When the probability that
conditional variance process in regime 2 equals 0.9, even in the stationary case we will no longer have consistent estimator of $\beta_1$. Since regime 2 is more volatile regime, here the high probability that the conditional variance is in such regime may be the reason that we fail to get consistent estimator. We also notice that the estimates of parameters in regime 1 turn out to have larger MSE. It confirms our assertion that the small probability in one regime affects the performance of estimates in that regime. Figure 6-8 provide the estimated density of MLE estimates summarized in table 2. The MLE estimates are approximately unbiased and consistent even the stationarity condition is violated. Figure 9-10 provide the estimated density of MLE estimates summarized in table 2.

4 Empirical Study

In this section we apply the threshold model to empirical data and find a good fit of the threshold GARCH model. We begin with a brief description of the data set and follow with the estimation of the threshold model. Then we discuss the estimation results.

4.1 The Data

The data set consists of 20 stocks in the major market index (MMI)\(^6\). We obtain the data of most stocks for the period from Jan. 2, 1970 to Dec. 31, 2008, except for AXP and T. We have data only start from May 18, 1977 for AXP and Jan 2, 1984 for T. We choose the stocks from MMI because they are well known and highly capitalized stocks representing a broad range of industries and they generally exhibit a high level of trading activity. Return data are obtained from daily stock file of the Center for Research in Security Prices (CRSP) and accessed from Wharton Research Data Services (WRDS).

\(^6\)The firms in the MMI are American Express (AXP), AT&T (T), Chevron (CHV), Coca-Cola (KO), Disney (DIS), Dow Chemical (DOW), Du Pont (DD), Eastman Kodak (EK), Exxon (XOM), General Electric (GE), General Motors (GM), International Business Machines (IBM), International Paper (IP), Johnson & Johnson (JNJ), McDonald's (MCD), Merck (MRK), 3M (MMM), Philip Morris (MO), Procter and Gamble (PG), and Sears (S).
The exogenous threshold variable we used in this empirical study is volatility index. The Chicago Board Options Exchange (CBOE) Volatility Index is a key measure of market expectations of near-term (30-day) volatility conveyed by S&P 500 stock index option prices. It is a weighted blend of prices for a range of options on the S&P 500 index. The volatility index is calculated and disseminated in real-time by CBOE. We obtain the data from CBOE website from Jan. 2 1990 to Nov. 18 2005. The volatility index provides an estimate of the implied volatility of the S&P 500 index over the next 30-days, therefore it is approximately independent of current volatility, however the expectation on the future market condition would definitely affect the market activity today.

The summary statistics for the returns is presented in Table 2. The columns report the sample minimum, maximum, mean, standard deviation, coefficient of skewness, and coefficient of kurtosis.

<table>
<thead>
<tr>
<th>Stocks</th>
<th># of observations</th>
<th>minimum</th>
<th>maximum</th>
<th>mean</th>
<th>std</th>
<th>skewness</th>
<th>kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>AXP</td>
<td>7890</td>
<td>-.2623</td>
<td>.1856</td>
<td>5.91e-04</td>
<td>.0215</td>
<td>-.1186</td>
<td>11.0912</td>
</tr>
<tr>
<td>T</td>
<td>6274</td>
<td>-.2194</td>
<td>.2025</td>
<td>6.04e-04</td>
<td>.0173</td>
<td>.1174</td>
<td>14.7662</td>
</tr>
<tr>
<td>CHV</td>
<td>9845</td>
<td>-.1667</td>
<td>.2085</td>
<td>6.44e-04</td>
<td>.0166</td>
<td>.2082</td>
<td>10.5984</td>
</tr>
<tr>
<td>KO</td>
<td>9845</td>
<td>-.2469</td>
<td>.1957</td>
<td>5.70e-04</td>
<td>.0160</td>
<td>-.0172</td>
<td>15.5269</td>
</tr>
<tr>
<td>DIS</td>
<td>9845</td>
<td>-.2909</td>
<td>.1910</td>
<td>6.24e-04</td>
<td>.0209</td>
<td>-.1985</td>
<td>13.1012</td>
</tr>
<tr>
<td>DOW</td>
<td>9845</td>
<td>-.1932</td>
<td>.1524</td>
<td>5.03e-04</td>
<td>.0181</td>
<td>-.1165</td>
<td>10.0208</td>
</tr>
<tr>
<td>DD</td>
<td>9845</td>
<td>-.1827</td>
<td>.1147</td>
<td>4.45e-04</td>
<td>.0166</td>
<td>-.0177</td>
<td>8.1179</td>
</tr>
<tr>
<td>EK</td>
<td>9844</td>
<td>-.3024</td>
<td>.2406</td>
<td>1.84e-04</td>
<td>.0191</td>
<td>-.2118</td>
<td>22.0107</td>
</tr>
<tr>
<td>XOM</td>
<td>9845</td>
<td>-.2343</td>
<td>.1790</td>
<td>6.77e-04</td>
<td>.0146</td>
<td>.0071</td>
<td>19.2241</td>
</tr>
<tr>
<td>GE</td>
<td>9845</td>
<td>-.1749</td>
<td>.1361</td>
<td>5.63e-04</td>
<td>.0164</td>
<td>.0250</td>
<td>9.8170</td>
</tr>
<tr>
<td>GM</td>
<td>9844</td>
<td>-.3111</td>
<td>.3511</td>
<td>2.25e-04</td>
<td>.0251</td>
<td>.4792</td>
<td>31.6393</td>
</tr>
<tr>
<td>IBM</td>
<td>9842</td>
<td>-.2296</td>
<td>.1316</td>
<td>4.02e-04</td>
<td>.0169</td>
<td>.0614</td>
<td>12.9150</td>
</tr>
<tr>
<td>IP</td>
<td>9845</td>
<td>-.2695</td>
<td>.1912</td>
<td>3.37e-04</td>
<td>.0186</td>
<td>-.1214</td>
<td>13.2700</td>
</tr>
<tr>
<td>JNJ</td>
<td>9845</td>
<td>-.1835</td>
<td>.1223</td>
<td>5.87e-04</td>
<td>.0155</td>
<td>-.0489</td>
<td>8.9948</td>
</tr>
<tr>
<td>MCD</td>
<td>9844</td>
<td>-.1942</td>
<td>.1622</td>
<td>7.35e-04</td>
<td>.0184</td>
<td>.0174</td>
<td>9.7235</td>
</tr>
<tr>
<td>MRK</td>
<td>9845</td>
<td>-.2678</td>
<td>.1303</td>
<td>5.49e-04</td>
<td>.0166</td>
<td>-.4530</td>
<td>14.5407</td>
</tr>
<tr>
<td>MMM</td>
<td>9845</td>
<td>-.2598</td>
<td>.1154</td>
<td>4.49e-04</td>
<td>.0148</td>
<td>-.3594</td>
<td>16.3442</td>
</tr>
<tr>
<td>MO</td>
<td>9843</td>
<td>-.2300</td>
<td>.1638</td>
<td>8.36e-04</td>
<td>.0174</td>
<td>-.1798</td>
<td>12.6358</td>
</tr>
<tr>
<td>PG</td>
<td>9845</td>
<td>-.3138</td>
<td>.2220</td>
<td>5.85e-04</td>
<td>.0148</td>
<td>-1.1529</td>
<td>42.4104</td>
</tr>
<tr>
<td>S</td>
<td>9845</td>
<td>-.2713</td>
<td>.2866</td>
<td>3.47e-04</td>
<td>.0221</td>
<td>-.1551</td>
<td>27.8540</td>
</tr>
</tbody>
</table>
4.2 Estimation of the Threshold GARCH model

In this section we apply the threshold GARCH model to the above data set which contains 20 stocks from MMI. We assume that the return series follows the threshold GARCH model where the volatility index is the exogenous trigger variable.

Since the return series in our threshold GARCH model is assumed to be a zero mean process, we first remove the mean from the returns. In addition to a constant we also filter the AR effect to order 5:

$$r_t = R_t - \mu - \sum_{j=1}^{5}\delta_j R_{t-j}$$

where $R_t$ is the observed return, $\mu$ is the mean, and $\delta_j$ is the coefficient of AR variables.

The volatility index is the exogenous trigger in this threshold GARCH model, so we also need to determine the threshold value for this variable. Now the model that needs to be estimated has a set of parameters as a function of threshold values.

$$r_t = \sigma_t \varepsilon_t$$

$$\begin{cases} 
\sigma_t^2 = \omega_0 + \alpha_0 r_{t-1}^2 + \beta_0 \sigma_{t-1}^2 & \text{if } y_{t-1} \leq y_0 \\
\sigma_t^2 = \omega_1 + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2 & \text{if } y_{t-1} > y_0 
\end{cases}$$

Define the log likelihood function as follows:

$$\ln L_T(\theta) = -\frac{1}{2} \sum_{t=1}^{T} \ln \sigma_t^2 - \frac{1}{2} \sum_{t=1}^{T} \frac{r_t^2}{\sigma_t^2}$$

where $\theta = \{\omega_0, \omega_1, \alpha_0, \alpha_1, \beta_0, \beta_1\}$. Then the MLE estimator $\hat{\theta}$ is a function of $y_0$:

$$\hat{\theta}(y_0) = arg\max T^{-1} \ln L$$

To estimate the threshold value we divide the sample of threshold variable into 40 intervals and the 39 grid points correspond to 2.5 percentile point to 97.5 percentile point. To evaluate the
model performance, we use the Akaike’s information criterion and Bayesian information criterion to compare the threshold GARCH model to standard GARCH (1,1) model.

\[
\begin{align*}
AIC &= 2k - 2lnL \\
BIC &= -2lnL + kln(n)
\end{align*}
\]

where \( n \) is the sample size and \( k \) is the number of parameters estimated in the model.

4.3 Preliminary Results

The MLE estimates are presented in Table 3.
Table 3: The Estimates of Threshold GARCH model Using VIX as Threshold Variable

<table>
<thead>
<tr>
<th></th>
<th>$y_0$</th>
<th>$\alpha_0$</th>
<th>$\alpha_1$</th>
<th>$\beta_0$</th>
<th>$\beta_1$</th>
<th>LLF</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>AXP</td>
<td>GARCH(1,1)</td>
<td>.062</td>
<td>.0107</td>
<td>.9368</td>
<td>.0007</td>
<td>12202</td>
<td>-24398</td>
<td>-24379</td>
</tr>
<tr>
<td>TS</td>
<td>$y_{90}$</td>
<td>.0414</td>
<td>.1476</td>
<td>.9552</td>
<td>.0006</td>
<td>12218</td>
<td>-24425</td>
<td>-24393</td>
</tr>
<tr>
<td>T</td>
<td>GARCH(1,1)</td>
<td>.0317</td>
<td>.0006</td>
<td>.9678</td>
<td>.0006</td>
<td>10400</td>
<td>-20794</td>
<td>-20775</td>
</tr>
<tr>
<td>TS</td>
<td>$y_{625}$</td>
<td>.0530</td>
<td>.0142</td>
<td>.9356</td>
<td>.0110</td>
<td>10419</td>
<td>-20828</td>
<td>-20797</td>
</tr>
<tr>
<td>CHV</td>
<td>GARCH(1,1)</td>
<td>.0611</td>
<td>.0091</td>
<td>.9209</td>
<td>.0118</td>
<td>13590</td>
<td>-27174</td>
<td>-27155</td>
</tr>
<tr>
<td>TS</td>
<td>$y_{60}$</td>
<td>.0354</td>
<td>.0860</td>
<td>.9148</td>
<td>.0207</td>
<td>13607</td>
<td>-27204</td>
<td>-27171</td>
</tr>
<tr>
<td>KO</td>
<td>GARCH(1,1)</td>
<td>.0485</td>
<td>.0065</td>
<td>.9490</td>
<td>.0077</td>
<td>13672</td>
<td>-27338</td>
<td>-27319</td>
</tr>
<tr>
<td>TS</td>
<td>$y_{925}$</td>
<td>.0272</td>
<td>.1378</td>
<td>.9704</td>
<td>.0863</td>
<td>13694</td>
<td>-27378</td>
<td>-27346</td>
</tr>
<tr>
<td>DIS</td>
<td>GARCH(1,1)</td>
<td>.0555</td>
<td>.0114</td>
<td>.9374</td>
<td>.0152</td>
<td>12437</td>
<td>-24867</td>
<td>-24848</td>
</tr>
<tr>
<td>TS</td>
<td>$y_{35}$</td>
<td>.0634</td>
<td>.1023</td>
<td>.7402</td>
<td>.0539</td>
<td>12476</td>
<td>-24943</td>
<td>-24911</td>
</tr>
<tr>
<td>DOW</td>
<td>GARCH(1,1)</td>
<td>.0575</td>
<td>.0080</td>
<td>.9632</td>
<td>.0083</td>
<td>12846</td>
<td>-25686</td>
<td>-25667</td>
</tr>
<tr>
<td>TS</td>
<td>$y_{675}$</td>
<td>.0523</td>
<td>.0620</td>
<td>.9260</td>
<td>.0132</td>
<td>12857</td>
<td>-25704</td>
<td>-25671</td>
</tr>
<tr>
<td>DD</td>
<td>GARCH(1,1)</td>
<td>.0343</td>
<td>.0051</td>
<td>.9632</td>
<td>.0058</td>
<td>12950</td>
<td>-25894</td>
<td>-25875</td>
</tr>
<tr>
<td>TS</td>
<td>$y_{925}$</td>
<td>.0247</td>
<td>.1186</td>
<td>.9723</td>
<td>.0539</td>
<td>12967</td>
<td>-25924</td>
<td>-25892</td>
</tr>
<tr>
<td>EK</td>
<td>GARCH(1,1)</td>
<td>.1990</td>
<td>.0005</td>
<td>.6231</td>
<td>.0002</td>
<td>11973</td>
<td>-23940</td>
<td>-23921</td>
</tr>
<tr>
<td>TS</td>
<td>$y_{625}$</td>
<td>.1290</td>
<td>.2969</td>
<td>.2063</td>
<td>.0002</td>
<td>12051</td>
<td>-24092</td>
<td>-24060</td>
</tr>
<tr>
<td>XOM</td>
<td>GARCH(1,1)</td>
<td>.0589</td>
<td>.0083</td>
<td>.9302</td>
<td>.0094</td>
<td>13768</td>
<td>-27531</td>
<td>-27511</td>
</tr>
<tr>
<td>TS</td>
<td>$y_{95}$</td>
<td>.0400</td>
<td>.1835</td>
<td>.9521</td>
<td>.0001</td>
<td>13787</td>
<td>-27563</td>
<td>-27531</td>
</tr>
<tr>
<td>GE</td>
<td>GARCH(1,1)</td>
<td>.047</td>
<td>.0004</td>
<td>.9522</td>
<td>.0002</td>
<td>13478</td>
<td>-26950</td>
<td>-26931</td>
</tr>
<tr>
<td>TS</td>
<td>$y_{775}$</td>
<td>.0179</td>
<td>.1019</td>
<td>.9774</td>
<td>.0173</td>
<td>13508</td>
<td>-27006</td>
<td>-26973</td>
</tr>
<tr>
<td>GM</td>
<td>GARCH(1,1)</td>
<td>.0611</td>
<td>.0093</td>
<td>.9268</td>
<td>.0114</td>
<td>11596</td>
<td>-23186</td>
<td>-23167</td>
</tr>
<tr>
<td>TS</td>
<td>$y_{675}$</td>
<td>.0352</td>
<td>.0916</td>
<td>.9454</td>
<td>.0120</td>
<td>11611</td>
<td>-23212</td>
<td>-23180</td>
</tr>
<tr>
<td>IBM</td>
<td>GARCH(1,1)</td>
<td>.0542</td>
<td>.0085</td>
<td>.9434</td>
<td>.0089</td>
<td>12641</td>
<td>-25277</td>
<td>-25257</td>
</tr>
<tr>
<td>TS</td>
<td>$y_{925}$</td>
<td>.0311</td>
<td>.2118</td>
<td>.9664</td>
<td>.0062</td>
<td>12662</td>
<td>-25314</td>
<td>-25282</td>
</tr>
<tr>
<td>IP</td>
<td>GARCH(1,1)</td>
<td>.0518</td>
<td>.0079</td>
<td>.9419</td>
<td>.0083</td>
<td>12643</td>
<td>-25281</td>
<td>-25261</td>
</tr>
<tr>
<td>TS</td>
<td>$y_{575}$</td>
<td>.0305</td>
<td>.1276</td>
<td>.8713</td>
<td>.0321</td>
<td>12680</td>
<td>-25350</td>
<td>-25318</td>
</tr>
</tbody>
</table>
The parameters in the threshold GARCH model are significant for 19 out of 20 stocks except $\beta_1$ for EK. The likelihood and the model selection criteria AIC and BIC suggest the better fit of data of the threshold GARCH model than traditional GARCH model. Since we use the volatility index as the threshold variable for all 20 stocks, it is possible that some stocks are more sensitive to the market conditions than others, therefore we observe in some stocks the estimated parameters are very similar in both regimes. For some stocks we have the sum of estimated parameters greater than 1 in one regime, but consider the probability that conditional variance stays in that regime, we will still have the stationary process. The probability is given by the location of the threshold value in the sample space of the threshold variable, it is very clear in our estimation results since we use 2.5 percentile as the increment. For example the estimated parameters in regime 2 have a sum of 1.0172, but the threshold value estimated is $y_{65}$, that means there is only 5% of chances that the conditional variance shifts to the regime 2, therefore the stationarity conditions still hold. Though this simple empirical exercise is not sufficient proof of the threshold GARCH model, it provides a perspective of the further exploration and application of such a simple but flexible model.
5 Conclusion

In this paper we define a threshold GARCH model with exogenous triggers to describe the conditional variance dynamics. This model can capture both the regime-switching feature and the effect of exogenous variables such as volatility index on the regime shifts of the conditional variance process. The model is flexible to accommodate other exogenous variables that can trigger the regime changes. Although the theoretical asymptotic distribution of the estimator is hard to establish, we show by simulation that the MLE estimator is approximately unbiased and consistent. The model is also applied to the data set with 20 stocks, the performance of the model is compared with standard GARCH (1,1) model and showed good fit of the data.
Appendix A

Proof of Proposition 1

We rewrite the conditional variance equation as:

\[ \sigma_t^2 = \omega_0 (1 - S_{t-1}) + \omega_1 S_{t-1} + (\alpha_0 (1 - S_{t-1}) + \alpha_1 S_{t-1}) r_{t-1}^2 + (\beta_0 (1 - S_{t-1}) + \beta_1 S_{t-1}) \sigma_{t-1}^2 \]

\[ = \omega_0 + (\omega_1 - \omega_0) S_{t-1} + [(\alpha_0 + (\alpha_1 - \alpha_0) S_{t-1}) \epsilon_{t-1}^2 + \beta_0 + (\beta_1 - \beta_0) S_{t-1}] \sigma_{t-1}^2 \]  \( (1) \)

Let

\[ c_0 = \omega_0 (1 - \pi) + \omega_1 \pi \quad c_1 = \omega_1 - \omega_0 \]
\[ a_0 = \alpha_0 (1 - \pi) + \alpha_1 \pi \quad a_1 = \alpha_1 - \alpha_0 \]
\[ b_0 = \beta_0 (1 - \pi) + \beta_1 \pi \quad b_1 = \beta_1 - \beta_0 \]
\[ B_t = S_t - \pi \]

Then (1) can be rewritten as:

\[ \sigma_t^2 = c_0 + c_1 B_{t-1} + [\epsilon_{t-1}^2 (a_0 + a_1 B_{t-1}) + b_0 + b_1 B_{t-1}] \sigma_{t-1}^2 \]  \( (2) \)

The transformation makes the random variables in the coefficient have mean zero. Back substitution in (2) results in:

\[ \sigma_t^2 = c_0 + c_1 B_{t-1} + \sum_{n=1}^{k-1} (c_0 + c_1 B_{t-1-n}) \prod_{m=1}^{n} [\epsilon_{t-m}^2 (a_0 + a_1 B_{t-m}) + b_0 + b_1 B_{t-m}] \]

\[ + \prod_{m=1}^{k} [\epsilon_{t-m}^2 (a_0 + a_1 B_{t-m}) + b_0 + b_1 B_{t-m}] \sigma_{t-k}^2 \]

Let we define \( S_n(t) \) as:

\[ S_n(t) = \prod_{m=1}^{n} [\epsilon_{t-m}^2 (a_0 + a_1 B_{t-m}) + b_0 + b_1 B_{t-m}] \]

26
Then we have \( \frac{1}{n} \ln(S_n(t)) \xrightarrow{a.s.} E(\ln[\varepsilon_{t-m}^2(a_0 + a_1 B_{t-m}) + b_0 + b_1 B_{t-m}]) \).

It’s very clear that if \( E(\ln[\varepsilon_{t-m}^2(a_0 + a_1 B_{t-m}) + b_0 + b_1 B_{t-m}]) < 0 \), that is \(([\alpha_0 + \beta_0](1 - \pi) + \alpha_1 + \beta_1)\pi < 1\), then the terms \( S_n(t) \) are geometrically bounded as \( n \) increases and equation (2) has the solution:

\[
\sigma_t^2 = c_0 + c_1 B_{t-1} + \sum_{n=1}^{\infty} (c_0 + c_1 B_{t-1-n})S_n(t)
\]

Therefore the return process has a stationary solution given by:

\[
r_t = \sqrt{c_0 + c_1 B_{t-1} + \sum_{n=1}^{\infty} (c_0 + c_1 B_{t-1-n})S_n(t) \varepsilon_t}.
\]

Proof of Proposition 2

Taking expectation on both side of equation (3) we have:

\[
E(\sigma_t^2) = E(c_0 + c_1 B_{t-1}) + \sum_{n=1}^{\infty} E(c_0 + c_1 B_{t-1-n})E(S_n(t))
\]

Given that \( \varepsilon_t \) and \( S_t \) are independent, we have:

\[
E(S_n(t)) = E[\prod_{m=1}^{n} \{\varepsilon_{t-m}^2(a_0 + a_1 B_{t-m}) + b_0 + b_1 B_{t-m}\}]
\]

\[
= \prod_{m=1}^{n} [E(\varepsilon_{t-m}^2(a_0 + a_1 B_{t-m})) + E(b_0 + b_1 B_{t-m})]
\]

\[
= (a_0 + b_0)^n
\]

Provided that \( a_0 + b_0 < 1 \), that is \( (\alpha_0 + \beta_0)(1 - \pi) + (\alpha_1 + \beta_1)\pi < 1 \), the equation (4) becomes:

27
\[ E(\sigma_t^2) = \sum_{n=1}^{\infty} E(c_0 + c_1 B_{t-1-n}^2) E(S_n(t)) \]
\[ = c_0 + c_0 \sum_{n=1}^{\infty} (a_0 + b_0)^n \]
\[ = c_0 \sum_{n=0}^{\infty} (a_0 + b_0)^n \]
\[ = \frac{c_0}{1 - (a_0 + b_0)} \]
\[ = \frac{\omega_0 (1 - \pi) + \omega_1 \pi}{1 - [(\alpha_0 + \beta_0)(1 - \pi) + (\alpha_1 + \beta_1)\pi]} \]

**Proof of Proposition 3**

To examine the stationarity conditions for higher order moments of return, we check the second moment of \( \sigma_t^2 \):

\[ E(\sigma_t^4) = E[(c_0 + c_1 B_{t-1}) + \sum_{n=1}^{\infty} (c_0 + c_1 B_{t-1-n}) S_n(t)]^2 \]
\[ = E(c_0 + c_1 B_{t-1})^2 + \sum_{n=1}^{\infty} E[(c_0 + c_1 B_{t-1-n}) S_n(t)]^2 \]
\[ = E(c_0 + c_1 B_{t-1})^2 + \sum_{n=1}^{\infty} E(S_n(t)) + 2c_0c_1 \sum_{n=1}^{\infty} E(B_{t-1-n} S_n(t)) + 2c_0c_1 \sum_{n=1}^{\infty} E(B_{t-1} S_n(t)) \]
\[ + 2c_1^2 \sum_{n=1}^{\infty} E(B_{t-1-n} S_n(t)) + 2c_0c_1 \sum_{n=1}^{\infty} E(B_{t-1-n} S_n(t))^2 \]

Note that
\[ E(B_{t-1}S_n(t)) = E\{B_{t-1}[\varepsilon_{t-1}^2(a_0 + a_1 B_{t-1}) + b_0 + b_1 B_{t-1}]S_n(t - 1)\} \]
\[ = (a_1 + b_1)\pi(1 - \pi)(a_0 + b_0)^{n-1} \]

\[ E[\sum_{n=1}^{\infty} (c_0 + c_1 B_{t-1-n})S_n(t)]^2 = E[c_0 \sum_{n=1}^{\infty} S_n(t) + c_1 \sum_{n=1}^{\infty} B_{t-1-n}S_n(t)]^2 \]
\[ = E[c_0^2(\sum_{n=1}^{\infty} S_n(t))^2] + E[2c_0 c_1 \sum_{n=1}^{\infty} S_n(t) \sum_{n=1}^{\infty} B_{t-1-n}S_n(t)] \]
\[ + E[c_1 \sum_{n=1}^{\infty} B_{t-1-n}S_n(t)]^2 \]

\[ E[c_0^2(\sum_{n=1}^{\infty} S_n(t))^2] = c_0^2 \sum_{n=1}^{\infty} E(S_n^2(t)) + 2 \sum_{n=1}^{\infty} \sum_{l=n+1}^{\infty} E(S_n(t)S_l(t)) \]
\[ E[2c_0 c_1 \sum_{n=1}^{\infty} S_n(t) \sum_{n=1}^{\infty} B_{t-1-n}S_n(t)] = 0 \]
\[ E[c_1 \sum_{n=1}^{\infty} B_{t-1-n}S_n(t)]^2 = c_1^2 [\pi(1 - \pi) \sum_{n=1}^{\infty} E(S_n^2(t)) \]
\[ + 2 \sum_{n=1}^{\infty} \sum_{l=n+1}^{\infty} E(B_{t-1-n}B_{t-1-l}S_n(t)S_l(t)) \]

where
Let \( A = 2a_0^2 + (a_0 + b_0)^2 + (2a_1^2 + (a_1 + b_1)^2)\pi (1 - \pi) \), then

\[
E(S_n^2(t)) = E \left[ \prod_{m=1}^{n} (\varepsilon_{t-m}^2(a_0 + a_1B_{t-m}) + b_0 + b_1B_{t-m})^2 \right] \\
= E \left[ \prod_{m=1}^{n} \left( \varepsilon_{t-m}^4(a_0 + a_1B_{t-m})^2 + 2\varepsilon_{t-m}^2(a_0 + a_1B_{t-m})(b_0 + b_1B_{t-m}) \\
+ (b_0 + b_1B_{t-m})^2 \right) \right] \\
= [3a_0^2 + a_1^2\pi (1 - \pi) + 2a_0b_0 + 2a_1b_1\pi (1 - \pi) + b_0^2 + b_1^2\pi (1 - \pi)]^n \\
= [2a_0^2 + (a_0 + b_0)^2 + (2a_1^2 + (a_1 + b_1)^2)\pi (1 - \pi)]^n
\]

Therefore:

\[
E(S_n(t)S_l(t)) = E \left[ \prod_{m=1}^{n} (\varepsilon_{t-m}^2(a_0 + a_1B_{t-m}) + b_0 + b_1B_{t-m})^2 \right] \\
\prod_{m=n+1}^{\infty} (\varepsilon_{t-m}^2(a_0 + a_1B_{t-m}) + b_0 + b_1B_{t-m}) \right] \\
= A^n(a_0 + b_0)^{l-n}
\]

Provided that

\[
A = 2a_0^2 + (a_0 + b_0)^2 + (2a_1^2 + (a_1 + b_1)^2)\pi (1 - \pi) \\
= (2a_0^2 + (\alpha_0 + \beta_0)^2)(1 - \pi) + (2a_1^2 + (\alpha_1 + \beta_1)^2)\pi < 1
\]

30
The above equations can be simplified as:

\[
E[c_0^2(\sum_{n=1}^{\infty} S_n(t))^2] = c_0^2(\sum_{n=1}^{\infty} A^n(1 + 2\sum_{j=1}^{\infty} (a_0 + b_0)^j))
\]

\[
= c_0^2 \frac{A(1 + (a_0 + b_0))}{(1 - A)(1 - (a_0 + b_0))}
\]

\[
E[c_1\sum_{n=1}^{\infty} B_{t-1-n}S_n(t)]^2 = c_1^2\pi(1 - \pi)\sum_{n=1}^{\infty} A^n
\]

\[
= c_1^2\pi(1 - \pi) \frac{A}{(1 - A)}
\]

Substitute back to the expression of \( E(\sigma_t^4) \), we have:

\[
E(\sigma_t^4) = c_0^2 + c_1^2\pi(1 - \pi) + 2c_0^2\sum_{n=1}^{\infty} (a_0 + b_0)^n + 2c_0c_1\pi(1 - \pi)(a_1 + b_1)\sum_{n=1}^{\infty} (a_0 + b_0)^{n-1}
\]

\[
+ c_0^2 \frac{A(1 + a_0 + b_0)}{(1 - A)(1 - (a_0 + b_0)) + c_1^2\pi(1 - \pi) \frac{A}{(1 - A)}
\]

\[
= c_0^2 + c_1^2\pi(1 - \pi) + 2c_0^2(a_0 + b_0) \frac{1}{1 - (a_0 + b_0)} + 2c_0c_1\pi(1 - \pi) \frac{a_1 + b_1}{1 - (a_0 + b_0)}
\]

\[
+ c_0^2 \frac{A(1 + a_0 + b_0)}{(1 - A)(1 - (a_0 + b_0)) + c_1^2\pi(1 - \pi) \frac{A}{(1 - A)}
\]

\[
= c_0^2 \frac{1 + a_0 + b_0}{(1 - A)(1 - (a_0 + b_0))} + c_1\pi(1 - \pi) \frac{c_1(1 - (a_0 + b_0)) + 2c_0(a_1 + b_1)(1 - A)}{(1 - A)(1 - (a_0 + b_0))}
\]

**Proof of Proposition 4**

We have \( \sigma^2 = E(r_t^2) = \frac{c_0}{1 - (a_0 + b_0)} \), subtract mean from equation (2), we get:
\[
s_i^2 = c_1 B_{t-1} + [\varepsilon_{t-1}^2 (a_0 + a_1 B_{t-1}) + b_0 + b_1 B_{t-1}] \sigma^2_{t-1} - \sigma^2 (a_0 + b_0)
\]
\[
r_i^2 = c_1 B_{t-1} + (a_0 + a_1 B_{t-1}) r_{t-1}^2 + (b_0 + b_1 B_{t-1}) \sigma^2_{t-1} - \sigma^2 (a_0 + b_0) - h_t + r_t^2
\]

The above expression of \( r_t^2 - \sigma^2 \) is very similar to the equation (3.6) in Ding and Granger (1996). Interestingly, the above expression represents \( r_t^2 - \sigma^2 \) as an random coefficient ARMA(1,1) process if we write the compound error as an MA(1) process; letting \( \sigma^2_t (\varepsilon_t^2 - 1) = v_t \):

\[
\sigma^2_t (\varepsilon_t^2 - 1) - (b_0 + b_1 B_{t-1}) \sigma^2_{t-1} (\varepsilon_{t-1}^2 - 1) = v_t - (b_0 + b_1 B_{t-1}) v_{t-1}
\]

with \( E(v_t) = 0 \) and \( E(v_t v_s) = 0 \) for all \( t \neq s \). Therefore, without any formal proof, the information already gives us some idea of the behaviors of the auto covariances.

Multiply both sides by \( (r_{t-1}^2 - \sigma^2) \):

\[
(r_t^2 - \sigma^2) (r_{t-1}^2 - \sigma^2) = [a_0 + b_0 + (a_1 + b_1) B_{t-1}] (r_{t-1}^2 - \sigma^2)^2 + (a_0 + a_1 + b_1 + c_1) B_{t-1} (r_{t-1}^2 - \sigma^2)
\]

Taking expectation on both sides

\[
E(r_t^2 - \sigma^2) (r_{t-1}^2 - \sigma^2) = (a_0 + b_0) E(r_{t-1}^2 - \sigma^2)^2 - b_0 E(r_{t-1}^2 - \sigma^2)^2 (r_{t-1}^2 - \sigma^2)
\]

\[
\gamma(1) = (a_0 + b_0) \gamma(0) - 2b_0 E(\sigma_{t-1}^4)
\]
where $\gamma(1) = E(r_t^2 - \sigma^2)(r_{t-1}^2 - \sigma^2)$ is the covariance between $r_t^2$ and $r_{t-1}^2$, and $\gamma(0) = E(r_{t-1}^2 - \sigma^2)^2$ is the variance of $r_{t-1}^2$. If we assume that the second moment of conditional variance or the fourth moment of residual exists, that is when the conditions in proposition 3 satisfied, then it can be shown that:

$$\rho(1) = \frac{c_0^2[2a_0 - 2a_0b_0A_0 + A_0(2a_1^2 + (a_1 + b_1)^2)\pi(1 - \pi)]}{c_0^2(2 + A - 3A_0^2) + 3c_1\pi(1 - \pi)[c_1(1 - A_0) + 2c_0(a_1 + b_1)(1 - A)]}$$

$$+ \frac{c_0^2(3a_0 + b_0(1 + 2A_0))\pi(1 - \pi)[c_1^2 + 2c_0c_1(a_1 + b_1)(1 - A)]}{c_0^2(2 + A - 3A_0^2) + 3c_1\pi(1 - \pi)[c_1(1 - A_0) + 2c_0(a_1 + b_1)(1 - A)]}$$

where $A_0 = a_0 + b_0$

It is easy to show that:

$$\gamma(k) = (a_0 + b_0)\gamma(k - 1) \text{ for } k \geq 2, \text{ and } \rho(k) = (a_0 + b_0)^{k-1}\rho(1)$$
### Appendix B

Table 5: The MLE estimates for parameters in Case 2

<table>
<thead>
<tr>
<th>Par</th>
<th>True Value</th>
<th>Estimate</th>
<th>MSE</th>
<th>True Value</th>
<th>Estimate</th>
<th>MSE</th>
<th>True Value</th>
<th>Estimate</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\omega_0$</td>
<td>0.02</td>
<td>0.0229</td>
<td>.0005</td>
<td>0.02</td>
<td>.0226</td>
<td>.0003</td>
<td>0.02</td>
<td>.0201</td>
<td>.0002</td>
</tr>
<tr>
<td>$\omega_1$</td>
<td>0.01</td>
<td>.0090</td>
<td>.0010</td>
<td>0.01</td>
<td>.0096</td>
<td>.0005</td>
<td>0.01</td>
<td>.0107</td>
<td>.0004</td>
</tr>
<tr>
<td>$\alpha_0$</td>
<td>0.25</td>
<td>.2342</td>
<td>.0192</td>
<td>0.25</td>
<td>.2383</td>
<td>.0081</td>
<td>0.25</td>
<td>.2243</td>
<td>.0111</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>0.25</td>
<td>.2415</td>
<td>.0313</td>
<td>0.25</td>
<td>.2370</td>
<td>.0123</td>
<td>0.25</td>
<td>.2265</td>
<td>.0183</td>
</tr>
<tr>
<td>$\beta_0$</td>
<td>0.5</td>
<td>.4768</td>
<td>.0649</td>
<td>0.5</td>
<td>.4880</td>
<td>.0159</td>
<td>0.5</td>
<td>.4546</td>
<td>.0291</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.75</td>
<td>.7973</td>
<td>.1577</td>
<td>0.9</td>
<td>.9314</td>
<td>.0439</td>
<td>1</td>
<td>.9020</td>
<td>.1091</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Par</th>
<th>True Value</th>
<th>Estimate</th>
<th>MSE</th>
<th>True Value</th>
<th>Estimate</th>
<th>MSE</th>
<th>True Value</th>
<th>Estimate</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\omega_0$</td>
<td>0.02</td>
<td>.0212</td>
<td>.0002</td>
<td>0.02</td>
<td>.0209</td>
<td>.0001</td>
<td>0.02</td>
<td>.0194</td>
<td>.0001</td>
</tr>
<tr>
<td>$\omega_1$</td>
<td>0.01</td>
<td>.0093</td>
<td>.0003</td>
<td>0.01</td>
<td>.0103</td>
<td>.0001</td>
<td>0.01</td>
<td>.0103</td>
<td>.0001</td>
</tr>
<tr>
<td>$\alpha_0$</td>
<td>0.25</td>
<td>.2482</td>
<td>.0092</td>
<td>0.25</td>
<td>.2463</td>
<td>.0024</td>
<td>0.25</td>
<td>.2307</td>
<td>.0074</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>0.25</td>
<td>.2477</td>
<td>.0096</td>
<td>0.25</td>
<td>.2483</td>
<td>.0044</td>
<td>0.25</td>
<td>.2364</td>
<td>.0096</td>
</tr>
<tr>
<td>$\beta_0$</td>
<td>0.5</td>
<td>.4811</td>
<td>.0375</td>
<td>0.5</td>
<td>.4949</td>
<td>.0042</td>
<td>0.5</td>
<td>.4611</td>
<td>.0235</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.75</td>
<td>.7740</td>
<td>.0509</td>
<td>0.9</td>
<td>.9063</td>
<td>.0085</td>
<td>1</td>
<td>.8962</td>
<td>.0914</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Par</th>
<th>True Value</th>
<th>Estimate</th>
<th>MSE</th>
<th>True Value</th>
<th>Estimate</th>
<th>MSE</th>
<th>True Value</th>
<th>Estimate</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\omega_0$</td>
<td>0.02</td>
<td>.0207</td>
<td>.0001</td>
<td>0.02</td>
<td>.0204</td>
<td>.0000</td>
<td>0.02</td>
<td>.0203</td>
<td>.0000</td>
</tr>
<tr>
<td>$\omega_1$</td>
<td>0.01</td>
<td>.0096</td>
<td>.0001</td>
<td>0.01</td>
<td>.0101</td>
<td>.0000</td>
<td>0.01</td>
<td>.0107</td>
<td>.0001</td>
</tr>
<tr>
<td>$\alpha_0$</td>
<td>0.25</td>
<td>.2487</td>
<td>.0023</td>
<td>0.25</td>
<td>.2482</td>
<td>.0011</td>
<td>0.25</td>
<td>.2567</td>
<td>.0015</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>0.25</td>
<td>.2473</td>
<td>.0030</td>
<td>0.25</td>
<td>.2477</td>
<td>.0018</td>
<td>0.25</td>
<td>.2599</td>
<td>.0027</td>
</tr>
<tr>
<td>$\beta_0$</td>
<td>0.5</td>
<td>.4897</td>
<td>.0090</td>
<td>0.5</td>
<td>.4987</td>
<td>.0018</td>
<td>0.5</td>
<td>.5050</td>
<td>.0021</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.75</td>
<td>.7651</td>
<td>.0152</td>
<td>0.9</td>
<td>.9012</td>
<td>.0033</td>
<td>1</td>
<td>.9754</td>
<td>.0061</td>
</tr>
</tbody>
</table>

34
Figure 6: Kernel Smoothing Density Estimates of MLE for Stationary Returns with Fourth Moment when $\pi = 0.5$
Figure 7: Kernel Smoothing Density Estimates of MLE for Stationary Returns without Fourth Moment when $\pi = 0.5$.
Figure 8: Kernel Smoothing Density Estimates of MLE for Non-Stationary Returns when $\pi = 0.5$.
<table>
<thead>
<tr>
<th>Par</th>
<th>True Value</th>
<th>Estimate</th>
<th>MSE</th>
<th>True Value</th>
<th>Estimate</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \omega_0 )</td>
<td>0.02</td>
<td>.0175</td>
<td>.0007</td>
<td>0.02</td>
<td>.0210</td>
<td>.0009</td>
</tr>
<tr>
<td>( \omega_1 )</td>
<td>0.01</td>
<td>.0118</td>
<td>.0007</td>
<td>0.01</td>
<td>.0123</td>
<td>.0010</td>
</tr>
<tr>
<td>( \alpha_0 )</td>
<td>0.25</td>
<td>.2361</td>
<td>.0257</td>
<td>0.25</td>
<td>.2353</td>
<td>.0215</td>
</tr>
<tr>
<td>( \alpha_1 )</td>
<td>0.25</td>
<td>.2437</td>
<td>.0304</td>
<td>0.25</td>
<td>.2412</td>
<td>.0236</td>
</tr>
<tr>
<td>( \beta_0 )</td>
<td>0.5</td>
<td>.5780</td>
<td>.0973</td>
<td>0.5</td>
<td>.5562</td>
<td>.0528</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>0.7</td>
<td>.7088</td>
<td>.0988</td>
<td>0.75</td>
<td>.7505</td>
<td>.0658</td>
</tr>
</tbody>
</table>

T=500

<table>
<thead>
<tr>
<th>Par</th>
<th>True Value</th>
<th>Estimate</th>
<th>MSE</th>
<th>True Value</th>
<th>Estimate</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \omega_0 )</td>
<td>0.02</td>
<td>.0179</td>
<td>.0003</td>
<td>0.02</td>
<td>.0185</td>
<td>.0004</td>
</tr>
<tr>
<td>( \omega_1 )</td>
<td>0.01</td>
<td>.0104</td>
<td>.0002</td>
<td>0.01</td>
<td>.0112</td>
<td>.0002</td>
</tr>
<tr>
<td>( \alpha_0 )</td>
<td>0.25</td>
<td>.2482</td>
<td>.0143</td>
<td>0.25</td>
<td>.2444</td>
<td>.0124</td>
</tr>
<tr>
<td>( \alpha_1 )</td>
<td>0.25</td>
<td>.2501</td>
<td>.0119</td>
<td>0.25</td>
<td>.2491</td>
<td>.0080</td>
</tr>
<tr>
<td>( \beta_0 )</td>
<td>0.5</td>
<td>.5543</td>
<td>.0472</td>
<td>0.5</td>
<td>.5452</td>
<td>.0349</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>0.7</td>
<td>.7075</td>
<td>.0352</td>
<td>0.75</td>
<td>.7425</td>
<td>.0200</td>
</tr>
</tbody>
</table>

T=1000

<table>
<thead>
<tr>
<th>Par</th>
<th>True Value</th>
<th>Estimate</th>
<th>MSE</th>
<th>True Value</th>
<th>Estimate</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \omega_0 )</td>
<td>0.02</td>
<td>.0180</td>
<td>.0002</td>
<td>0.02</td>
<td>.019</td>
<td>.0002</td>
</tr>
<tr>
<td>( \omega_1 )</td>
<td>0.01</td>
<td>.0099</td>
<td>.0001</td>
<td>0.01</td>
<td>.0103</td>
<td>.0001</td>
</tr>
<tr>
<td>( \alpha_0 )</td>
<td>0.25</td>
<td>.2510</td>
<td>.0069</td>
<td>0.25</td>
<td>.2488</td>
<td>.0051</td>
</tr>
<tr>
<td>( \alpha_1 )</td>
<td>0.25</td>
<td>.2489</td>
<td>.0040</td>
<td>0.25</td>
<td>.2475</td>
<td>.0019</td>
</tr>
<tr>
<td>( \beta_0 )</td>
<td>0.5</td>
<td>.5434</td>
<td>.0264</td>
<td>0.5</td>
<td>.5251</td>
<td>.0156</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>0.7</td>
<td>.7086</td>
<td>.0151</td>
<td>0.75</td>
<td>.7503</td>
<td>.0060</td>
</tr>
</tbody>
</table>

T=2000
Figure 9: Kernel Smoothing Density Estimates of MLE for Stationary Returns with Fourth Moment when $\pi = 0.9$.
Figure 10: Kernel Smoothing Density Estimates of MLE for Stationary Returns without Fourth Moment when $\pi = 0.9$
References


