Estimation of Volatility on the Small Sample with Generalized Maximum Entropy

Quanrui Song[1(✉)], Songsak Sriboonchitta[1,2], Somsak Chanaim[1,2], and Chongkolnee Rungruang[3]

[1] Faculty of Economics, Chiang Mai University, Chiang Mai 50200, Thailand
song_quanrui@cmu.ac.th
[2] Center of Excellence in Econometrics, Chiang Mai University, Chiang Mai 50200, Thailand
[3] Faculty of Commerce and Management, Prince of Songkla University, Trang Campus, Trang 92000, Thailand

Abstract. Generalized autoregressive conditional heteroscedasticity (GARCH) provides useful techniques for modeling the dynamic volatility model. Several estimation techniques have been developed over the years, for examples Maximum likelihood, Bayesian, and Entropy. Among these, entropy can be considered an efficient tool for estimating GARCH model since it does not require any distribution assumptions which must be given in Maximum likelihood and Bayesian estimators. Moreover, we address the problem of estimating GARCH model characterized by ill-posed features. We introduce a GARCH framework based on the Generalized Maximum Entropy (GME) estimation method. Finally, in order to better highlight some characteristics of the proposed method, we perform a Monte Carlo experiment and we analyze a real case study. The results show that entropy estimator is successful in estimating the parameters in GARCH model and the estimated parameters are close to the true values.

Keywords: Volatility · GARCH(1, 1) model
Generalized Maximum Entropy

1 Introduction

Estimation of volatility is very important in financial economics, because volatility is a measure of uncertainty on observed time series of financial data such as stock price or stock index. Many studies realized that the volatility of financial data should not be constant overtime but invariably varying through time. The popular model to estimate the time-varying volatility is Autoregressive conditional heteroskedasticity model (ARCH) proposed by Engel [1] in 1982 and was extended by Bollerslev [2]. This paper introduces a new volatility model called Generalized autoregressive conditional heteroskedasticity model (GARCH(p,q)) in the following form

$$\varepsilon_t = \sigma_t \nu_t, \ , \nu_t \sim N(0,1), t = 1, 2 \cdots , T,$$

V.-N. Huynh et al. (Eds.): IUKM 2018, LNAI 10758, pp. 324–334, 2018.
https://doi.org/10.1007/978-3-319-75429-1_27

where

$$\sigma_t^2 = \omega_0 + \sum_{i=1}^{p} \omega_{1i}\sigma_{t-i}^2 + \sum_{j=1}^{q} \omega_{2j}\varepsilon_{t-j}^2,\ \omega_0, \omega_{1i}, \omega_j > 0\ \forall i = 1, \cdots, p,\ j = 1, \cdots, q.$$

In this study, we consider GARCH(1,1) model because this is an extension of ARCH model and relies only on past observation and on past volatility. In general, GARCH parameters have been estimated by using Maximum Likelihood (MLE) approach which assumes normality. However, the assumption of conditional normality is not always appropriate. Maximum Entropy (ML) modeling which has a flexible functional form to use with many distributions has been applied in financial field. Park and Bera [3] applied two separate maximum entropy densites in ARCH model (MEARCH model) where moment functions are selected based on the sample to estimate NYSE stock returns. They show the MEARCH model was useful to capture the behavior of the sample.

From assumption on long-term stability, GARCH model will give a wrong answer if data is not stability trend. Financial time series data generally are characterized with a large sample size and structural change. To be consistent with this stability assumption, we suggest small sample size for estimating GARCH(1,1) model. Hwang and Valls Pereira [4] investigated ML estimation with small sample in GARCH model with non-negative Bollerslev's condition that guarantees positive conditional volatility and they showed GARCH models with small sample problems from their results that the estimated parameters are negatively biased. They also suggested the minimum size of sample needed for GARCH(1,1) model to be 500 observations.

When the data has a heavy-tailed distribution, the analysis of GARCH time series data by using quasi maximum likelihood estimation (QMLE) can lead to inconsistency in parameter noted by Lee et al. [5]. Who applied maximum entropy to estimate GARCH(1,1) model for 503 observations of S&P 500 index.

Generalized Maximum Entropy (GME) method for ARCH model can be found in the book by Golan et al. [6]. In this paper, we use GME to estimate GARCH(1,1) parameters because this method does not require the large observation and assumption about distribution function for innovation. This method has only independent assumption for all random variables and support space for each random variable. We show the result by using simulation and applying the model to estimate volatility on stock price returns with small number of observations.

2 Methodology

Let $\varepsilon_t,\ t = 1, \cdots, T$ is sequence random variable from time series data with mean zero. The GARCH(1,1) model is defined by

$$\varepsilon_t = \sigma_t \cdot \nu_t,\ t = 1, 2, \cdots, T, \quad (1)$$

$$\sigma_t^2 = \omega_0 + \omega_1\sigma_{t-1}^2 + \omega_2\varepsilon_{t-1}^2,\ t = 1, 2, \cdots, T, \quad (2)$$

where $\omega_0 > 0, \omega_1, \omega_2 \in (0,1)$ and stationary condition is $\omega_1 + \omega_2 < 1$ and $\nu_t \sim F(\cdot)$ is sequence of independent random variable or innovation. In this study, the parameters $\omega_0, \omega_1, \omega_2$ are estimated using generalized maximum entropy (GME). The basic idea of this estimator is that the entropy, which refers to an amount of the uncertainty, is maximized subject to model and data constraints. Here, we consider the Shannon's entropy measure proposed by Shannon [7]. This Shannon's entropy is represented by the amount of the uncertainty of a discrete probability distribution and the sum of all outcomes probability equals to one. The constraint primal problem for ARCH model can be written as follows:

$$\max H(P_0, P_1, P_2, W_1, W_2, \cdots, W_T) = H(P_0) + H(P_1) + H(P_2) + \sum_{t=1}^{T} H(W_t), \quad (3)$$

subject to

$$\sqrt{\left(\sum_{i=1}^{k} z_{0i} p_{0i}\right)} \left(\sum_{i=1}^{k} s_i w_{1i}\right) = \varepsilon_1,$$

$$\sqrt{\sum_{i=1}^{k} z_{0i} p_{0i} + \left(\sum_{i=1}^{k} z_{1i} p_{1i}\right) \left(\frac{\varepsilon_1}{\sum_{i=1}^{k} s_i w_{1i}}\right)^2 + \left(\sum_{i=1}^{k} z_{2i} p_{2i}\right) \cdot \varepsilon_1^2} \left(\sum_{i=1}^{k} s_i w_{2i}\right) = \varepsilon_2,$$

$$\sqrt{\sum_{i=1}^{k} z_{0i} p_{0i} + \left(\sum_{i=1}^{k} z_{1i} p_{1i}\right) \left(\frac{\varepsilon_2}{\sum_{i=1}^{k} s_i w_{2i}}\right)^2 + \left(\sum_{i=1}^{k} z_{2i} p_{2i}\right) \cdot \varepsilon_2^2} \left(\sum_{i=1}^{k} s_i w_{2i}\right) = \varepsilon_3,$$

$$\vdots$$

$$\sqrt{\sum_{i=1}^{k} z_{0i} p_{0i} + \left(\sum_{i=1}^{k} z_{1i} p_{1i}\right) \left(\frac{\varepsilon_{T-1}}{\sum_{i=1}^{k} s_i w_{(T-1)i}}\right)^2 + \left(\sum_{i=1}^{k} z_{2i} p_{2i}\right) \cdot \varepsilon_{T-1}^2} \left(\sum_{i=1}^{k} s_i w_{(T-1)i}\right) = \varepsilon_T,$$

$$\frac{\sum_{i=1}^{k} z_{0i} p_{0i}}{1 - \sum_{i=1}^{k} z_{1i} p_{1i} - \sum_{i=1}^{k} z_{1i} p_{0i}} = var(\varepsilon_t),$$

$$\sum_{i=1}^{k} p_{0i}, \sum_{i=1}^{k} p_{1i}, \sum_{i=1}^{k} p_{2i}, \sum_{i=1}^{k} w_{1i}, \cdots = \sum_{i=1}^{k} w_{Ti} = 1,$$

$$p_{ji}, w_{ti} \in (0,1) \; \forall j = 0, 1, 2, \; i = 1, \cdots, k, \; t = 1, \cdots, T,$$

where $H(P_j) = -\sum_{i=1}^{k} p_{ji}\log(p_{ji}), j = 0,1,2$ and $H(W_j) = -\sum_{i=1}^{k} w_{ti}\log(w_{ti}),\ t = 1,\cdots,T$, $z_{0i}, z_{1i}, z_{2i}, s_i$ are the discrete support space, and $var(\varepsilon_t)$ is the sample variance. After optimizing this function, we can estimate $\widehat{\omega_0}, \widehat{\omega_1}, \widehat{\omega_2}, \widehat{\nu_t}$ by

$$\widehat{\omega}_0 = \sum_{i=1}^{k} z_{0i}p_{0i},\ \widehat{\omega}_1 = \sum_{i=1}^{k} z_{1i}p_{1i},\ \widehat{\omega}_2 = \sum_{i=1}^{k} z_{2i}p_{2i},\ \widehat{\nu}_t = \sum_{i=1}^{k} s_i w_{ti},\ t = 1,2,\cdots,T.$$

The standard deviation of parameters $\widehat{\omega_0}, \widehat{\omega_1}$ and $\widehat{\omega_2}$ is to be estimated by

$$\text{std. of } \widehat{\omega}_j = \sqrt{\sum_{i=1}^{k}(z_{ji} - \widehat{\omega}_j)^2 p_{ji}},\ j = 0,1,2.$$

3 Simulation Study

In this section, a simulation study was conducted to evaluate performance and accuracy of GME estimation in GARCH(1,1) model with small observation $T = \{50, 100\}$. For every support space, we define 5 points support for $\omega_0, \omega_1, \omega_2$ with $z_0 = z_1 = z_2 = [0, 0.25, 0.50, 0.75, 1]$ and support space for innovation ν_t is $[-10, -5, 0, 5, 10]$ for all $t = 1, 2, \cdots, T$. We simulated the data from the GARCH(1,1) model with 1,000 paths, where the innovation is assumed to have standard normal distribution $N(0, 1)$. We consider $\omega_0, \omega_1, \omega_2$ for 5 cases.

The simulation results are provided in Tables 1, 2, 3, 4 and 5 and the similar results are obtained. According to Table 1, by the simulation we observe that the estimation mean of ω_0 is underestimated but those for ω_1, ω_2 are overestimated for both T = 50 and 100. From Table 2, we see that the estimation mean of ω_0 and ω_1 are underestimated while ω_2 value is overestimated for both T = 50 and 100. From Table 3, the different results are obtained, the means of ω_0 and ω_1 are overestimated but that for ω_2 is underestimated for both T = 50 and 100. From Table 4, we see that the means of ω_0, ω_1 and ω_2 are

Table 1. Case 1: $\omega_0 = \omega_1 = \omega_2 = 0.2$

	True	T = 50				T = 100			
		Mean	Std	Quantile 5%	Quantile 95%	Mean	Std	Quantile 5%	Quantile 95%
ω_0	0.2	0.0393	0.0215	0.0168	0.0769	0.0428	0.0202	0.0241	0.0712
ω_0	0.2	0.3533	0.0181	0.3202	0.3804	0.3468	0.0159	0.3180	0.3715
ω_0	0.2	0.3298	0.0133	0.3077	0.3505	0.3067	0.0107	0.2893	0.3228
Entropy		83.3738	0.1663	83.1607	83.6627	163.3534	0.1802	163.1272	163.6612

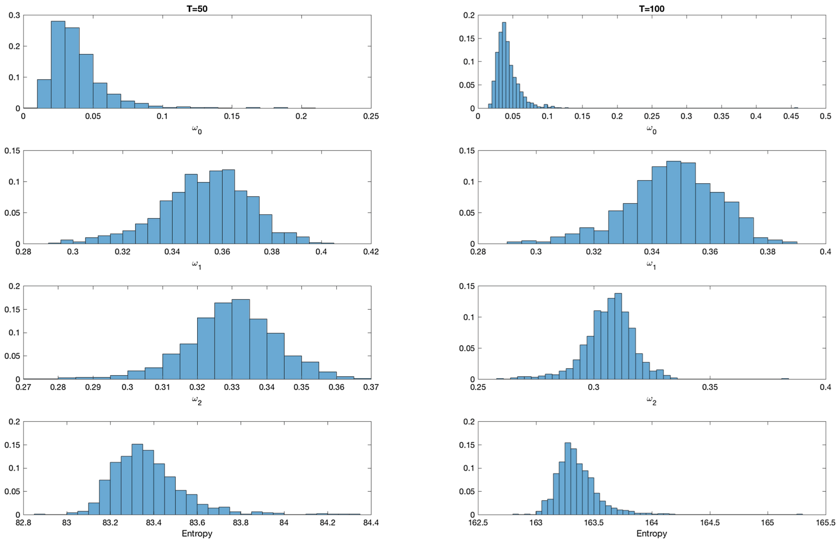

Fig. 1. Histogram of parameter estimates from simulation $T = 50$(left), $T = 100$(right)

Table 2. Case 2: $\omega_0 = .7,\ \omega_1 = .5,\ \omega_2 = 0.1$

	True	T = 50				T = 100			
		Mean	Std	Quantile 5%	Quantile 95%	Mean	Std	Quantile 5%	Quantile 95%
ω_0	0.7	0.5099	0.0287	0.4624	0.5569	0.5318	0.0228	0.4985	0.5708
ω_1	0.5	0.4363	0.0255	0.3907	0.4760	0.4605	0.0215	0.4251	0.4962
ω_2	0.1	0.3862	0.0236	0.3446	0.4235	0.3617	0.0190	0.3292	0.3930
Entropy		84.5511	0.0928	84.3883	84.6699	164.4394	0.0982	164.2645	164.5694

Table 3. Case 3: $\omega_0 = .2,\ \omega_1 = .2,\ \omega_2 = 0.7$

	True	T = 50				T = 100			
		Mean	Std	Quantile 5%	Quantile 95%	Mean	Std	Quantile 5%	Quantile 95%
ω_0	0.2	0.4048	0.1826	0.0975	0.6389	0.4405	0.1771	0.1000	0.6342
ω_1	0.2	0.3741	0.0669	0.2938	0.4846	0.3912	0.0804	0.2920	0.4946
ω_2	0.7	0.3777	0.0737	0.2920	0.5061	0.3909	0.0768	0.2898	0.5190
Entropy		84.3664	0.5714	83.5372	84.8916	164.4499	0.9400	162.1926	165.1179

overestimated for both T = 50 and 100; however, they are close to the true values. Finally, from Table 5, the estimation mean of ω_0 is underestimated and overestimated for T = 50 and T = 100 respectively; the estimation mean of ω_1 is underestimated for both T = 50 and T = 100; and the mean of ω_2 is overestimated for both T = 50 and T = 100.

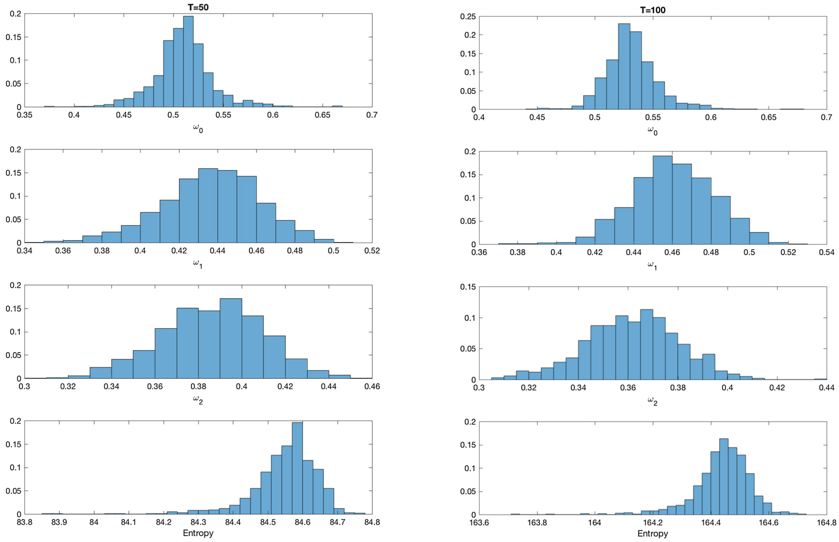

Fig. 2. Histogram of parameter estimates from simulation $T = 50$(left), $T = 100$(right)

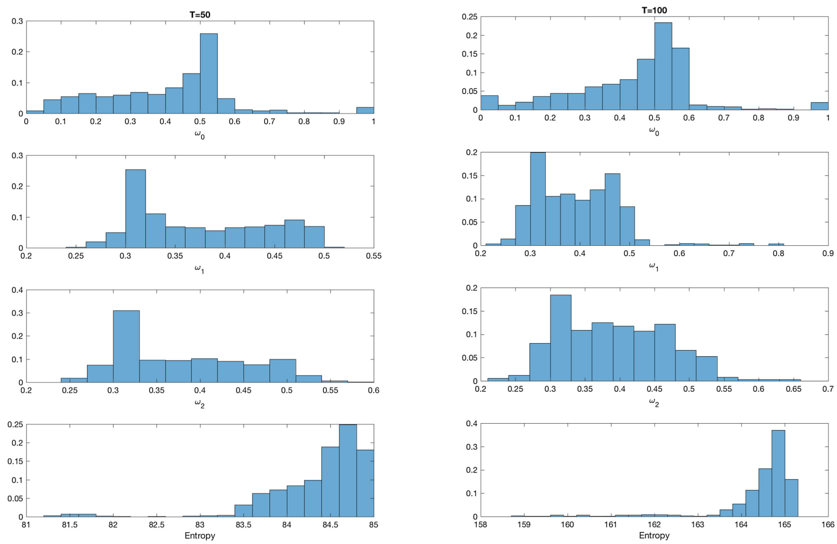

Fig. 3. Histogram of parameter estimates from $T = 50$(left), $T = 100$(right)

Table 4. Case 4: $\omega_0 = .5,\ \omega_1 = .4,\ \omega_2 = 0.4$

	True	T = 50				T = 100			
		Mean	Std	Quantile 5%	Quantile 95%	Mean	Std	Quantile 5%	Quantile 95%
ω_0	0.5	0.5365	0.0809	0.4402	0.6538	0.5389	0.1380	0.4393	0.6611
ω_1	0.4	0.4495	0.0393	0.3691	0.4975	0.4800	0.0518	0.4124	0.5203
ω_2	0.4	0.4290	0.0435	0.3457	0.4904	0.4275	0.0408	0.3620	0.4882
Entropy		84.5230	0.4421	84.0229	84.7723	164.3091	1.0100	161.6720	164.7983

Source calculation.

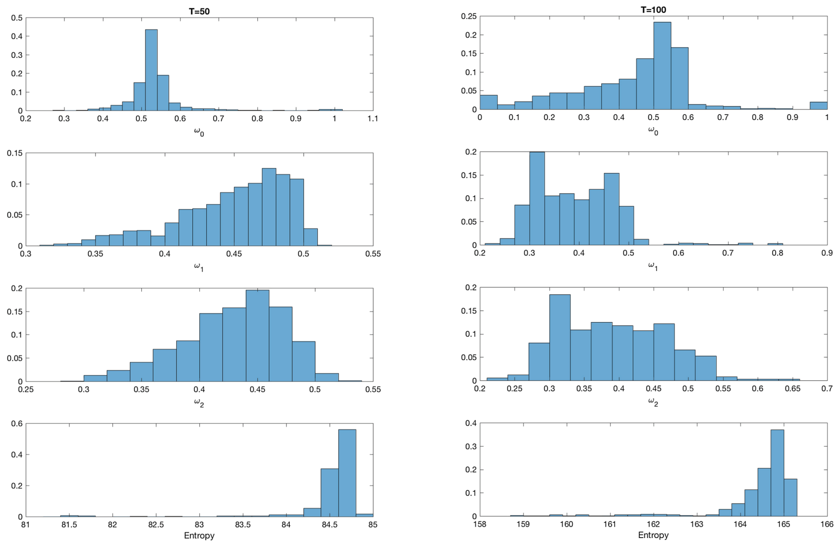

Fig. 4. Histogram of parameter estimates from $T = 50$(left), $T = 100$(right)

Table 5. Case 5: $\omega_0 = .6,\ \omega_1 = .7,\ \omega_2 = 0.2$

	True	T = 50				T = 100			
		Mean	Std	Quantile 5%	Quantile 95%	Mean	Std	Quantile 5%	Quantile 95%
ω_0	0.6	0.6476	0.1462	0.5285	0.9912	0.4819	0.3302	0.0006	0.9904
ω_1	0.7	0.5170	0.0202	0.4921	0.5442	0.5833	0.0849	0.5118	0.7500
ω_2	0.2	0.4565	0.0231	0.4227	0.4862	0.4569	0.0809	0.3933	0.5686
Entropy		83.77	0.9866	81.3374	84.5333	162.4112	1.9288	159.2213	163.3010

The overall results of the simulation are likely to perform well for all case studies as the estimated mean parameters are not far away from the true values. Moreover, we also plot the histogram of estimated parameters from 1,000 paths. We present all case studies and plot in Figs. 1, 2, 3, 4 and 5. We can observe that most estimated parameters are close to the true values and that the standard

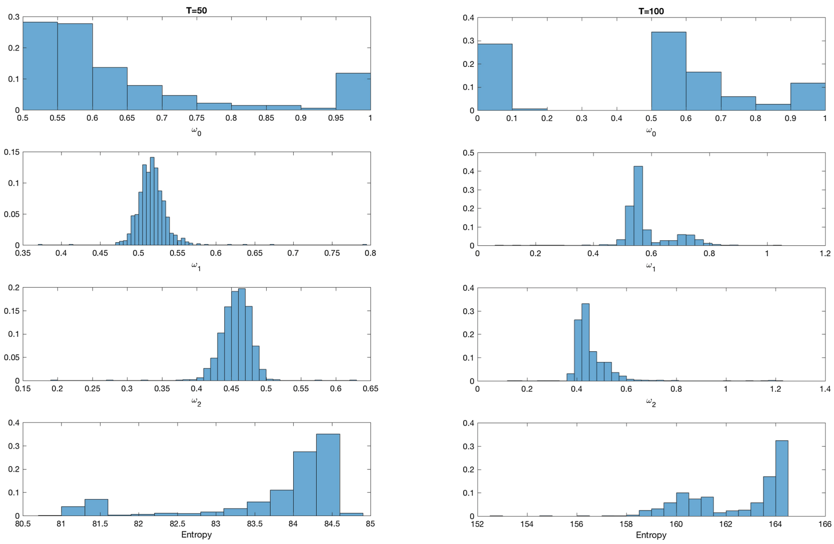

Fig. 5. Histogram of parameter estimates from $T = 50$(left), $T = 100$(right)

deviations are rather small. We may thus conclude that the proposed method estimates the GARCH(1,1) parameters quite well.

4 Real Data Application

In this section, we compare the performance of the entropy GARCH(1,1) using the stock index data. We consider closing stock price and daily return from Advanced Micro Devices, Inc. (AMD) from March 1, 2017 to July 28, 2017. The data is obtained from Thomson Reuters DataStream. We plot the daily closing price and return of AMD in Fig. 6. The summary statistics is shown in Table 6.

In this application study, we consider the simple GARCH(1,1) to estimate the volatility of the AMD return. The support spaces of the GARCH parameters are specified just like in the simulation study section. The estimated results are shown in Table 7. We can see that our estimated standard errors of $\omega_0, \omega_1, \omega_2$ in Table 7 are large and all parameters are insignificant. We suspect that our support spaces are perhaps too large and thereby leading a high standard error. Therefore, we try to estimate again but using a narrower range of support spaces for $\omega_0, \omega_1, \omega_2$. We define a new support space as $z_0 = z_1 = z_2 = [0, 0.125, 0.25, 0.375, 0.5]$ and s_i as $[-10, -5, 0, 5, 10]$. The results are shown in Table 8, and when compared to Table 7, values of parameter estimates change and standard errors decrease; however, we can still obtain the significant results. We also plot the conditional variance and innovation error from our GARCH(1,1) (Fig. 7).

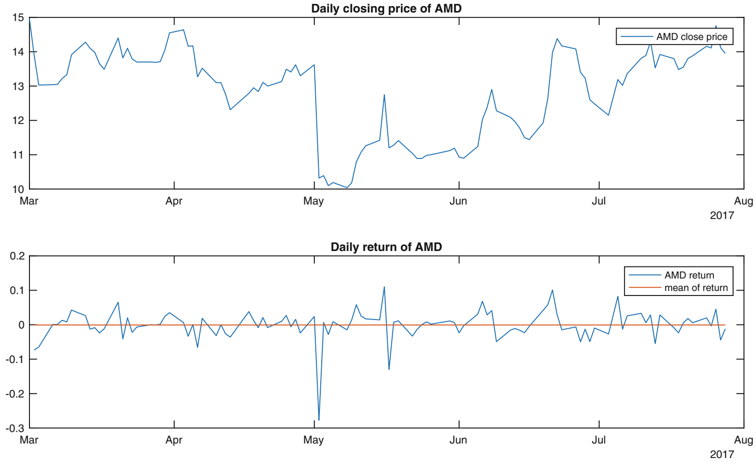

Fig. 6. Closing stock price (up) and return of AMD (down)

Table 6. Summary statistics of AMD

Mean	Median	Std	Min	Max	Skewness	Kurtosis	Obs.
−0.0007	0.0008	0.0443	−0.2775	0.1102	−2.2995	16.958	104

Table 7. Estimation of GARCH(1,1) parameter by GME

Parameter	Value	Std
ω_0	7.3761×10^{-4}	0.0137
ω_1	0.3689	0.3375
ω_2	0.2496	0.2927
Entropy	168.8636	

The support space for every parameter is the same as in the simulation part.

Table 8. Parameters estimated from GME with the support space

Parameter	Value	Std
ω_0	0.0011	0.0118
ω_1	0.2333	0.1763
ω_2	0.2079	0.1735
Entropy	169.3612	

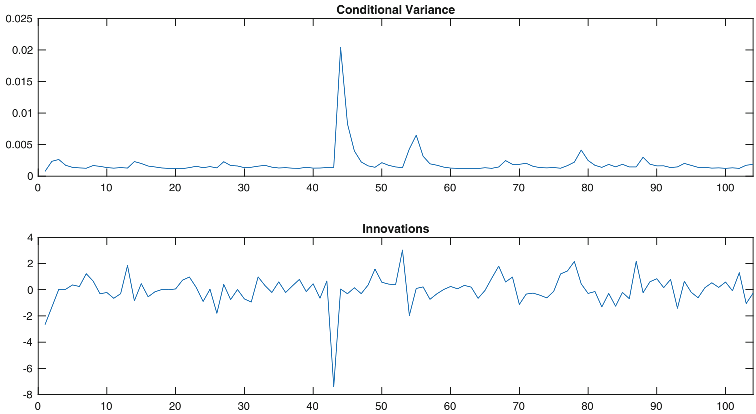

Fig. 7. Estimation of conditional variance (upper) and innovation (lower)

5 Conclusions and Future Research

It is not easy to get the big data with certain economic situation or stable environment. Thus many estimations face the ill-posed problem. In this study, GME estimator is proposed to estimate the unknown parameters in GARCH model. The simulation results show that the GME is workable well on some values of parameters since either underestimated or overestimated results are obtained in some parameters. However, the results are still acceptable in this study. From the real data analysis, we cannot find the significant result of the GARCH parameters. The problem of GME estimation in GARCH(1,1) is that the value of parameter estimates depends on the support space. We find that the narrow support space will lead a smaller standard error of the parameters.

Therefore in future research, we should find the new method to estimate a GARCH model or other volatility models that can handle the small observation problem. GME method should also be extended to GARCH(p,q) and other stochastic volatility (IGRACH, GJR-GRACH, NGARCH etc.) models.

Acknowledgement. The authors thank Mr. Woraphon Yamaka for his suggestions for GME method and applied to the GARCH(1,1) model. We would like to thank the referee for giving comments on the manuscript.

References

1. Engle, R.F.: Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom Inflation. Econometrica **50**(4), 987–1007 (1982)
2. Bollerslev, T.: Generalized autoregressive conditional heteroskedasticity. J. Econometrics **31**(3), 307–327 (1986)
3. Park, S.Y., Bera, A.K.: Maximum entropy autoregressive conditional heteroskedasticity model. J. Econometrics **150**(2), 219–230 (2009)

4. Hwang, S., Valls Pereira, P.L.: Small sample properties of GARCH estimates and persistence. Eur. J. Finance **12**(6–7), 473–494 (2006)
5. Lee, J., Lee, S., Park, S.: Maximum entropy test for GARCH models. Stat. Methodol. **22**, 8–16 (2015)
6. Golan, A., Judge, G., Miller, D.: Maximum Entropy Econometrics: Robust Estimation with Limited Data. Wiley, New York (1997)
7. Shannon, C.E.: A mathematical theory of communication. Bell Syst. Tech. J. **27**, 379–423 (1948)