Estimating and forecasting instantaneous volatility through a duration model: An assessment based on VaR [*][†]

Takayuki Morimoto[‡][§]

Graduate School of Social Sciences, Hiroshima University

2004 Far Eastern Meeting of the Econometric Society
July 2, 2004, Yonsei University, Seoul, Korea

Abstract

In order to forecast one-step ahead volatility, we calculated jump intensity by using estimated parameters of a duration model of price change. In this procedure, we do not assume any distribution on log-return. Although we do not make any distributional assumption, we may practically choose a suitable distribution e.g. Normal, student, etc, including empirical density, when we calculate a VaR (Value at Risk) with an instantaneous volatility to check the prediction performance. Furthermore, we compare the goodness of fit among assumed distributions of log-return. We find that fat tail distributions such as NIG, Laplace, are well fitted to the actual high frequency data listed on the Tokyo stock exchange 1st section from 4 Jan. 2001 to 28 June 2001.

1   Introduction

In the recent decade, the rapid development of computer technologies render the new type data known as high frequency data easily available. This kind of data contains many information in intraday trades which may perform an important role in the market microstructure analysis, and has some characteristics which are not found in a daily or weekly data. They are as follows. First, a trade interval (hereafter duration) is not constant or occurs non-equidistantly. Second, a periodicity called intraday seasonality exists within a day. Thus we cannot analyze such a data with usual time series modeling.

[‡]e-mail: morimot@hiroshima-u.ac.jp

[§]URI: http://home.hiroshima-u.ac.jp/rcfe/

And a study of pricing derivatives and a study of risk management are two underpinning in investment science. Although GARCH and SV models are standard model to analyze volatility in usual time series framework in this field, we cannot directly use those models to analyze high frequency data because of its peculiar characteristics.

Engle and Russell (1998) proposed a new model to deal with such a time series. We apply their framework to model a stochastic time interval as a duration, and derive a instantaneous volatility using a jump intensity. In order to give a verification of forecasting performance of the model, we use the (stochastically occurred) one-step ahead VaR calculated from the percentile of an arbitrary distribution fitted to actual log-return as in Giot (2000, 2002). Furthermore we compare the goodness of fit of among assumed distributions to actual log-return using the previous result.

This paper consists of three sections. In Section 1, we introduce the theoretical background for price duration, duration model, instantaneous volatility, seasonal adjustment, (stochastically occurred) one-step ahead VaR. In Section 2, using the high frequency data listed on the Tokyo stock exchange 1st section, we estimate the parameter of a duration model by MLE, and calculate an instantaneous volatility and (stochastically occurred) one-step ahead VaR. In order to verify the forecasting performance, likelihood test is executed. In the final Section, according to the result of examining the forecasting performance of (stochastically occurred) one-step ahead VaR, we compare the goodness of fit among assumed distributions of log-return.

2 Theoretical background

2.1 Price duration

Let S_i be a logarithmic stock price, and $r_i = S_i - S_{i-1}$ be its rate of return, where subscript i denotes a i th trade, $i = 1, 2, \dots$. If a cumulative sum of absolute value of r_i, i.e., $|r_i|$, overshoots a predefined constant threshold c_p, we regard that a jump has occurred. The time of a jump point is denoted by T_j, where $j = 1, 2, \dots$. The figure 1 shows this procedure. We redefine price duration $d_j = T_j - T_{j-1}$ by T_j .
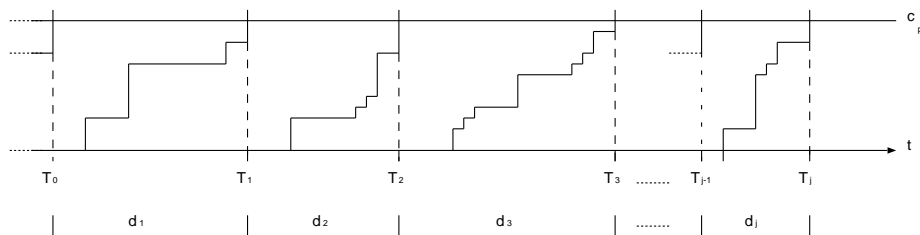

Figure 1: Price duration

2.2　Duration models

We define T_j, $j = 0, 1, 2, \ldots$ by a time when a jump occurs. T_j is a stochastic process occurring at an irregular time span. The time duration between j and $j - 1$ is defined by

$$d_j = T_j - T_{j-1}.$$

Sometimes it is assumed that

$$d_j = \psi_j \varepsilon_j \quad \varepsilon_j \sim \text{i.i.d. Non-negative R.V.},$$

where $\psi_j = \mathrm{E}[d_j | \mathcal{F}_{j-1}]$ and $\mathcal{F}_{j-1}$ is information sets until $j - 1$. Furthermore, analogous to GARCH, Engle and Russell (1998) advocates the ACD (p, q) (Autoregressive Conditional Duration) model as

$$\psi_j = \omega + \sum_{k=1}^{p} \alpha_k d_{j-k} + \sum_{k=1}^{q} \beta_k \psi_{j-k},$$

where ω_k , α_k , and β_k are constants satisfying $\omega > 0$, $\alpha, \beta \geq 0$. For guaranteeing the stationarity of the model, we impose the following constraint condition on parameters,

$$\sum_{k=1}^{p} \alpha_k + \sum_{k=1}^{q} \beta_k < 1.$$

Alternatively, it is assumed that

$$d_j = \exp(\Psi_j)\epsilon_j \quad \text{or} \quad \epsilon_j = \frac{d_j}{\exp(\Psi_j)},$$

then we have the Logarithmic-ACD (p, q) model,

$$\Psi_j = \omega + \sum_{k=1}^{p} \alpha_k \log(d_{j-k}) + \sum_{k=1}^{q} \beta_k \Psi_{j-k}. \tag{1}$$

See Bauwens and Giot (1999).

　　Although there are many candidates of distributions of ε_j such as Exponential, Gamma, and other non-negative distributions, among others we assume $\varepsilon_j \sim$ i.i.d. Weibull which is in the wider class than the exponential. We use the logarithmic-ACD model, because there are no restrictions on the sign of the parameters for positivity of $\exp(\Psi_j)\epsilon_j$. This model is called the Log-Weibull-ACD model. By definiton the density function of Weibull is

$$f(d_j) = \frac{\gamma}{d_j} \left(\frac{d_j \Gamma(1 + 1/\gamma)}{e^{\Psi_j}} \right)^{\gamma} e^{-\left(\frac{d_j \Gamma(1+1/\gamma)}{e^{\Psi_j}} \right)^{\gamma}},$$

where γ denotes shape parameter and Γ denotes gamma function. Thus the log likelihood function of this model is given by

$$\sum_{j=1}^{N} \left(\log(\gamma) - \log(d_j) + \gamma \log[d_j \Gamma(1 + 1/\gamma)] - \gamma \Psi_j - \left(\frac{d_j \Gamma(1 + 1/\gamma)}{e^{\Psi_j}} \right)^{\gamma} \right).$$

2.3 Instantaneous volatility

First we define a conditional intensity process as

$$\lambda(T|N(T), T_1, \ldots, T_{N(T)}) = \lim_{\Delta T \to 0} \frac{\Pr(N(T + \Delta T) > N(T)|N(T), T_1, \ldots, T_{N(T)})}{\Delta T},$$

where $\Pr(A|B)$ is a usual conditional probability, and $N(T)$ is the number of events that have occurred by time T. Let the instantaneous volatility be defined as

$$\hat{\sigma}^2(T) = \lim_{\Delta T \to 0} \mathrm{E} \left\{ \frac{1}{\Delta T} \left[\frac{S(T + \Delta T) - S(T)}{S(T)} \right]^2 \right\},$$

where $S(T)$ is a stock price associated with arriving time T, and taking limits, we obtain the conditional instantaneous volatility as

$$\hat{\sigma}^2(T|T_{N(T)}, \ldots, T_1) = \left(\frac{c_p}{S_{T_{N(T)}}} \right)^2 \lambda(T|T_{N(T)}, \ldots, T_1),$$

where c_p is an arbitrary prescribed constant.

We estimate the intensity through using Logarithmic-ACD model,

$$\Psi_j = \omega + \sum_{k=1}^{p} \alpha_k \log(d_{j-k}) + \sum_{k=1}^{q} \beta_k \Psi_{j-k}, \tag{2}$$

with

$$d_j = \exp(\Psi_j)\epsilon_j \quad \epsilon_j \sim \text{i.i.d. Non-negative R.V.},$$

where $\psi_j = \mathrm{E}[d_j|\mathcal{F}_{-1}]$. This is called the Logarithmic-ACD (p, q) model (see Bauwens and Giot (1999)).

As conditional intensity λ of Log-ACD model reduces to

$$\frac{1}{e^{\Psi}},$$

(see Giot (2000)), we obtain the instantaneous volatility of Log-ACD version is

$$\sigma^2(T|\mathcal{F}_{j-1}) = \left(\frac{c_p}{S(T_{j-1})} \right)^2 \frac{1}{e^{\Psi_j}}. \tag{3}$$

2.4 Adjusting intraday seasonality

In general, the duration of high frequency data has so-called intra-day seasonality, especially in Japanese stock market, there is a lunch time break, so the plot of its duration shows M-shaped. Using such data without any pretreatment we may not obtain precise estimation. Therefore we adjust the seasonality in the following way.

Let $\phi(t_i)$ be a deterministic term and d_i be non-adjusted duration series d_i, and we define a new series $d_j^\star$ as

$$d_j^\star = \frac{d_j}{\phi(t_j)}. \tag{4}$$

where $d_j^\star$ is called seasonal adjusted duration and $\phi(t_j)$ is called a time of day function. The $\phi(t_j)$ can be estimated by kernel or spline smoothing. Finally we obtain the seasonal adjusted volatility,

$$(\hat{\sigma}^\star)^2(T|\mathcal{F}_{j-1}) = \left(\frac{c_p}{S(T_{j-1})}\right)^2 \frac{1}{e^{\Psi_j^\star} \phi(t_j - 1)}. \tag{5}$$

2.5 (stochastically occurred) one-step ahead VaR

We can forecast (stochastically occurred) one-step ahead VaR using seasonal adjusted instantaneous volatility calculated above. The limit of percentile $1-\alpha$ is given by

$$\mathrm{VaR}(j) = z_{1-\alpha}\sqrt{(\hat{\sigma}^\star)^2(j|\mathcal{F}_{j-1})} * \sqrt{M * 60}, \tag{6}$$

where M denotes a time interval, $\sqrt{M * 60}$ have the unit of time interval change a minute into a second.

For obtaining percentile $z_{1-\alpha}$, we fit some suitable distribution to actual log-return data and estimate the parameters of those distributions. Four distributions that we use are as follows.

Normal distribution: It is the most fundamental distribution for investment science, but it is well known that Normal distribution may not describe the fat tail of financial time series. The density function is given by

$$f_{Normal}(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}, \quad -\infty < x < \infty.$$

Student distribution: Student distribution has fatter tail than Normal distribution. The density function with degree of freedom m is given by

$$f_{Student}(x) = \frac{\Gamma\left(\frac{m+1}{2}\right)}{\sqrt{m\pi}\Gamma\left(\frac{m}{2}\right)\left(1 + \frac{x^2}{m}\right)^{\frac{m+1}{2}}}, \quad -\infty < x < \infty.$$

If degree of freedom is larger than 30, then the distribution is almost the same as the normal distribution.

Laplace (double exponential) distribution: The density function of Laplace (double exponential) distribution with scale parameter β and location parameter μ is given by

$$f_{Laplace}(x) = \frac{1}{2\beta} \exp\left\{-\frac{|x-\mu|}{\beta}\right\}, \quad -\infty < x < \infty.$$

This one is obtained by the difference of two independent identical exponential distribution. See Abramowitz and Stegun 1972, p. 930.

Normal Inverse Gaussian (NIG) distribution: Following Barndorff-Nielsen (1997), Normal Inverse Gaussian (NIG) distribution is given by

$$f_{NIG}(x) = a(\alpha, \beta, \mu, \delta)q\left(\frac{x-\mu}{\delta}\right)^{-1} K_1\left\{\delta\alpha q\left(\frac{x-\mu}{\delta}\right)\right\}\exp(\beta x),$$

where

$$a(\alpha, \beta, \mu, \delta) = \pi^{-1}\alpha\exp(\delta\sqrt{(\alpha^2-\beta^2)} - \beta\mu),$$

$$q(x) = \sqrt{1+x^2},$$

and K_1 is the modified Bessel function of the third kind with coefficient 1. The parameter α, β, μ and δ are satisfied the condition $0 \leq |\beta| \leq \alpha$, $\mu \in R$ and $0 < \delta$.

3 Verification of forecasting performance

In this section, we verify forecasting performance of (stochastically occurred) one-step ahead VaR with the framework introduced in previous section. The data set which we use is the high frequency data of the following companies listed on the Tokyo stock exchange 1st section from 4 Jan. 2001 to 28 Sep. 2001. They are Nippon Steel(5401), HITACHI(6501), SONY(6578) and TOYOTA(7203). For later use, the original data is split into an estimation period (Jan. to Mar. 2001) and forecast period (Apr. to sep. 2001).

3.1 Estimation

First we estimate the parameter of LWACD model from actual data, and calculate the instantaneous volatility. We choose ten suitable threshold c_p for each data and distributions, and simulate price duration processes by using each c_p. For the seasonal adjustment, intraday time function $\hat{\phi}$ is estimated by Friedman's super smoother (see Friedman (1984a, 1984b)). After the adjustment LWACD $(1,1)$ model on the estimation period is given by

$$d_j^\star = \exp(\hat{\Psi}_j^\star)\epsilon_j, \quad \epsilon_j \sim \text{i.i.d. Weibull}(1, \hat{\gamma}),$$
$$\hat{\Psi}_j^\star = \hat{\omega} + \hat{\alpha}\log(d_{j-1}^\star) + \hat{\beta}\hat{\Psi}_{j-1}, \tag{7}$$

We estimate parameters, $\hat{\omega}(c_p)$, $\hat{\alpha}(c_p)$, $\hat{\beta}(c_p)$ and $\hat{\gamma}(c_p)$ by MLE, where the argument c_p in parenthesis indicates that these estimates depend on c_p . And we substitute $\hat{\Psi}_j^\star(c_p)$ into (6), then we can obtain seasonal adjusted instantaneous volatility $(\hat{\sigma}^\star(c_p))^2$. Hereafter for notational convenience the argument c_p is omitted.

3.2 Forecasting

One-step ahead VaR is calculated by substituting $(\hat{\sigma}^\star)^2$ for each c_p into

$$\text{VaR}(j) = z_{1-\alpha}\sqrt{(\hat{\sigma}^\star)^2(t|\mathcal{F}_{j-1})} * \sqrt{M * 60}, \tag{8}$$

and we verify the forecasting performance.

Since the VaR calculated by (8) is derived from stochastic duration $d_j^\star$ in (7), it occurred basically at a time different from actual data. Thus it is meaningless to discuss the forecasting performance by comparing values at different time. Therefore following Giot, we divide time interval into $M = 10, 15$ or 30 , regard the closing price in each time interval as a observed value for actual data, while regard the arithmetic average in each time interval as a observed value for estimated VaR. The reason why we take $M = 10, 15$ or 30 is that if we choose longer time interval like as 1 hour, then the sample size is too small because of opening time of TSE is only 4.5 hour within a day and that if we choose shorter time interval like as 5 minutes, then the possibility of no-trade in certain time interval becomes larger.

A percentile $z_{1-\alpha}$ in (8) is calculated by numerical integrations which are executed recursively for each parameters of Normal, Student, NIG, Laplace and Empirical. These parameters are estimated by MLE. Figure 4 shows an example. The solid line is actual data, and the dotted line is the VaR calculated by (8), on the forecasting period.

3.3 Likelihood Ratio Test

The failure rate of risk is defined as a ratio of the number of times that actual data exceeds VaR. Thus if the failure rate of risk is the nearer to the predefined percentile $1 - \alpha$, the forecasting performance of VaR is the better. Table 1 shows an example. We can say that if the entry of the table is the closer to nominal ratios in the head of each column, the VaR has the better forecasting performance. The outcome is classified into two cases depending on whether the VaR exceeds the actual data or not. Thus we can regard the outcome as Bernoulli trial. Then following Kupiec (1995), we can verify the forecasting performance of the VaR.
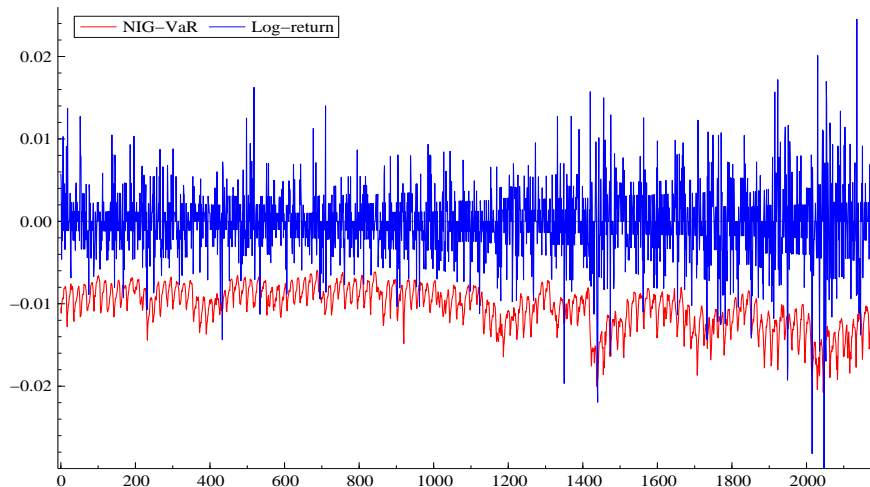
Figure 2: SONY (6758) sampled every 15min. (Apr. 2 - Sep. 28, 2001)

First, N denotes the number of total trade, and N_{ex} denotes the number that actual data exceeds to a limit of VaR within a forecasting period. Let p be the probability of the failure rate of risk, then N_{ex} follows

$$\Pr\{\nu\} = \binom{N}{\nu} p^\nu (1-p)^{N-\nu}, \quad \nu = 0, 1, ..., N.$$

Let define the likelihood of N_{ex} as

$$L(N_{ex}) = \frac{p_0^{N_{ex}}(1-p_0)^{N-N_{ex}}}{\hat{p}^{N_{ex}}(1-\hat{p})^{N-N_{ex}}}.$$

By setting $p_0 = 1 - \alpha$ and $\hat{p} = N_{ex}/N$ you can calculate the likelihood, and it is well-known that $2 \log L(N_{ex})$ follows $\chi^2(1)$. Hence we can execute hypothesis test with null hypothesis $H_0 : p = p_0$. Table 2 shows an example of log-likelihood ratio statistics $2 \log L(N_{ex})$ for each percentile $1 - \alpha$ of a VaR. If the statistic is not rejected, or accepted, then it is favorable to the assumed distribution. Asterisk "*" denotes the score defined by the following way: log-likelihood ratio statistic is not rejected at 90% then 10 points are assigned. Similarly at 75% 9 points, ..., 0.1% 1 point are assigned. Note that the score denotes only an order, the value itself is meaningless. Thus we can say that an entry of the table has more *, the density is more suitable for a return.

4 Comparing goodness of fit to distribution

Finally we compare a goodness of fit for each series by following steps.

8

Table 1: Failure Rate of Risk : SONY (6758) sampled every 15min.

	0.050	0.025	0.010	0.005	0.001
Normal	0.039778	0.020352	0.010638	0.006938	0.003700
Student	0.047032	0.021005	0.010502	0.006393	0.002740
NIG	0.056699	0.025606	0.010059	0.005944	0.000914
Laplace	0.059627	0.025489	0.010014	0.005007	0.000910
Empirical	0.049060	0.020174	0.011004	0.006419	0.002751

Table 2: Likelihood Ratio Test : SONY (6758) sampled every 15min.

	0.050	0.025	0.010	0.005	0.001
Normal	5.0989 *3	2.0445 *6	0.0872 *9	1.4555 *6	9.2743 *1
Student	0.4140 *8	1.5153 *6	0.0549 *9	0.7844 *7	4.4809 *4
NIG	1.9840 *6	0.0327 *9	0.0008 *10	0.3695 *8	0.0165 *9
Laplace	4.0484 *4	0.0214 *9	0.0000 *10	0.0000 *10	0.0182 *9
Empirical	0.0408 *9	2.2287 *6	0.2151 *8	0.8098 *7	4.5124 *4

1. Check the normality for actual data.

2. Choose 10 thresholds c_p 's for each time interval $M = 10, 15$ or 30 .

3. Execute the likelihood ratio test for VaR processes, 5 kinds of percentile $1 - \alpha$: 0.050, 0.025, 0.010, 0.005 and 0.001, 10 kinds of thresholds and 3 kinds of time intervals.

4. Score the results above by regarding the "goodness of forecasting performance of VaR" as the "goodness of fit to a distribution".

5. Execute step 1 to 4 for each actual data and distribution.

6. Rank the goodness of fit.

4.1 Testing normality

Before the study of a goodness of fit, we check the normality for actual data by Anderson-Darling test. This test may catch a fat tail better than Kolmogolv-Smirnov test. The Anderson-Darling test statistics A^2 is defined by

$$A^2 = -N - S,$$
$$S = \frac{1}{N} \cdot \sum_{i=1}^{N} (2i - 1) \left[\log \Phi_i + \log(1 - \Phi_{N+1-i}) \right],$$

where Φ is cumulative density function of standard normal distribution and N is a sample size. The null hypothesis H_0 : "data follows Normal distribution"

is rejected at significance level 10%, if $A^2 > 0.753$, at significance level 5% $A^2 > 1.159$. Table 3 shows the test result. This table reads the null hypothesis is rejected for all data sets, i.e. all data does not follow Normal distribution. Furthermore it is turned out that when the time interval becomes larger, the normality becomes greater for all data sets.

Table 3: The Anderson-Darling test statistics

	Nippon Steel	HITACHI	SONY	TOYOTA
10 min.	146.7047	20.2217	24.0964	97.7665
15 min.	84.2962	14.5411	20.7676	55.1298
30 min.	29.8662	6.2688	9.577	21.1466

4.2   Comparing goodness of fit

Finally we verify a goodness of fit to distribution. Table 4, 5 and 6 show the summary of the results of likelihood test for each 3 kinds of time interval $M = 10, 15$ or 30 , each 10 kinds of thresholds and 5 kinds of percentile of VaR. And table 7 shows the sums of result for each 3 kinds of time interval. In this table the top grade is

$$(\text{the points}) \times (\text{the number of time intervals}$$
$$\times \text{ the number of percentile of VaR}$$
$$\times \text{ the number of thresholds}).$$

i.e. $10 \times (3 \times 5 \times 10) = 1500$ is the highest score. Thus in this table we can read that NIG distribution is the fittest except for TOYOTA. Note that we are not taking into consideration the empirical density in this summary because it is not a probability distribution.

5   Concluding remarks

In this study, we estimate the instantaneous volatility through a duration model. And using it we calculate (stochastically occurred) one-step ahead VaR. And we investigate which density is the fittest to log-return by using the VaR as a measure of fitness. It seems that normal density is not suitable for financial time series such as log-return with fat tail and asymmetry. Finally, we obtain interesting result of the high frequency data of Tokyo stock exchange, which is natural and acceptable. There's a possibility that the method can be used for risk management of downside risk because it is only focusing on tail behavior.

References

[1] Abramowitz, M., and Stegun, I. (1972): *Handbook of Mathematical Functions*, Dover, New York.

[2] Barndorff-Nielsen O.E. (1997): "Normal inverse gaussian distributions and stochastic volatility modelling", *Scandinavian Journal of Statistics*, **24**, 1-13.

[3] Bauwens L. and P. Giot (1999): "The logarithmic ACD model : an application to the bid-ask quote process of two NYSE stocks", DP # 9789, CORE.

[4] Engle R. and J. Russell (1998): "Autoregressive conditional duration : a new model for irregularly-spaced transaction data", *Econometrica*, **66**, 1127-1162.

[5] Friedman J.H. (1984a): "SMART User's Guide. Laboratory for Computational Statistics", Stanford University Technical Report No. 1.

[6] Friedman J.H. (1984b): "A variable span scatterplot smoother. Laboratory for Computational Statistics", Stanford University Technical Report No. 5.

[7] Giot P. (2000): "Time transformations, intraday data and volatility models", *Journal of Computational Finance*, **4**, 31-62.

[8] Giot P. (2002): "Market risk models for intraday data", Discussion Paper 30, University of Namur.

[9] Kupiec P. (1995): "Techniques for verifying the accuracy of risk measurement models", *Journal of Derivatives*, **2**, 173-184.

Table 4: The score summary of likelihood ratio test results [10 min.]

	Normal	Student	NIG	Laplace	Empirical	The fittest
Nippon Steel	187	185	205	165	186	NIG
HITACHI	283	364	384	370	350	NIG
SONY	249	260	287	255	226	NIG
TOYOTA	194	206	230	210	191	NIG

Table 5: The score summary of likelihood ratio test results [15 min.]

	Normal	Student	NIG	Laplace	Empirical	The fittest
Nippon Steel	185	179	189	194	185	Laplace
HITACHI	368	341	346	306	393	Normal
SONY	233	270	288	257	284	NIG
TOYOTA	237	238	256	277	232	Laplace

Table 6: The score summary of likelihood ratio test results [30 min.]

	Normal	Student	NIG	Laplace	Empirical	The fittest
Nippon Steel	228	241	244	201	245	NIG
HITACHI	406	366	384	292	381	Normal
SONY	317	322	324	288	269	NIG
TOYOTA	276	240	258	271	234	Normal

Table 7: The score summary of likelihood ratio test results [10, 15, 30 min.]

	Normal	Student	NIG	Laplace	Empirical	The fittest
Nippon Steel	600	605	638	560	616	NIG
HITACHI	1057	1071	1114	968	1124	NIG
SONY	799	852	899	800	779	NIG
TOYOTA	707	684	744	758	657	Laplace