THE PREDICTABILITY OF OVERNIGHT INFORMATION

ZHUO ZHONG

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR

THE DEGREE OF MASTER OF SCIENCE IN FINANCE

SINGAPORE MANAGEMENT UNIVERSITY

2007

ABSTRACT

By decomposing close to close returns into close to open returns (overnight returns) and open to close returns (daytime returns), we test the predictability of overnight information, which is captured by absolute values of close to open returns, on daytime return volatility. Applying the stochastic volatility model, we find that overnight price changes contain important information to predict daytime volatility. The predictive power is highest at market opening and declines gradually over the trading day. Moreover, the predictive power is higher for inactive traded stocks than for actively traded stocks.

Contents

List of Tables

List of Figures

Acknowledgements

I would like to sincerely thank my supervisor Prof. Chunchi Wu, who has guided my thesis with reasonable strictness and inspiring suggestions. More importantly, besides the precious knowledge and research experience that he imparted to me, the right attitude to research he helped me to form and the passion he rekindled are really treasures in my future academic career.

Once more, Professor Wu, thank you for being such a great mentor and professor!

I would like to express my deep gratitude to my home institute, the Wang Yanan Institute for Studies in Economics (WISE), for continuing support for my endeavors to improve and to build up my academic work.

Special thanks go to Professors Yongmiao Hong, Hai Lin and Jun Tu for very helpful advices.

Also, many thanks to my classmates at SMU, Yaru Chen, Hongchao Han, Yan Teik Lee, Yong Tan. You all have made my one-year journey an exciting one. Perhaps, the best treasure one could always carry along is friendship.

At last, I would like to thank my parents and my beloved wife Ting Zhou for their strong support. Thanks for being such a great inspiration that keeps me going!

Dedicated to my beloved wife Ting Zhou and my parents!

1 Introduction

Conventionally, returns are defined as changes in closing price. The underlying assumption is that returns during the non-trading period (e.g., close-to-open) follow the same data generating process as the returns during the trading period (e.g., open-to-close). This assumption is arguable since there is usually no trade in the overnight period. Since information is incorporated into prices through trades, the data generating process in the overnight period is unlikely to be the same as that in the daytime trading period. Several studies have shown that the two data generating processes can be quite different. For example, Fama (1965), French and Roll (1986), Lockwood and Linn (1990) show that volatility over the trading period is significantly different from volatility over the non-trading period. Oldfield and Rogalski (1980), Hong and Wang (2000) have provided theoretical models to explain why they are different.

Market microstructure theory suggests that information revealed through trades affects security returns [see Hasbrouck (1991), Dufour and Engle (2000)] and volatility [see Xu, Chen, and Wu (2006)]. When the market is closed, no trading takes place and investors are prevented from adjusting their expectation based on trading information. This in turn can affect the return generating process. Craig, Dravid, and Richardson (1995), Chan, Chockalingam, and Lai (2000), Hong, Gallo, and Lee (2001) examine stock returns during the trading and non-trading periods. Barclay and Hendershott (2003) examine price discovery in after-hours trading period. These studies focus on the behavior of stock returns. Studies on volatility are relatively few. One example is Gallo and Pacini (1998) who employ the GARCH(1,1) model to incorporate the overnight information in forecasting daytime volatility.

This paper attempts to further explore the ability of overnight return to predict of daytime return volatility. Our approach differs from Gallo and Pacini (1998) in several ways. First, we employ the stochastic volatility model, which is considered to be more efficient than GARCH(1,1) model [see Kim, Shephard, and Chib (1998), Granger and Poon (2003)]. Second, our data spans over a much longer period. Third, we examine the relationship between overnight return and daytime volatility condition on the trading frequency.

Our paper contributes to the literature in several aspects. First, our results show that overnight returns contain important information for daytime return volatility. Second, we find that overnight return has higher predictive power for daytime return volatility for inactive trading stocks. This finding is consistent with the model of Easley et al. (1996) which predicts that the information content of trades is higher for infrequently traded stocks. Third, we find that overnight return of inactive stocks can predict the daytime return volatility over a longer horizon, which implies that overnight information plays a more important role in determining the volatility of inactive stocks. Finally, consistent with the finding of Kim, Shephard, and Chib (1998), Granger and Poon (2003), our paper show that the stochastic volatility model outperform the GARCH(1,1) model.

The rest of this paper is organized as follows. In Section 2, we review the literature. In Section 3, we describe the design of our empirical tests and discuss the Bayesian analysis. The data used in our empirical analysis is illustrated in Section 4.1. In Section 4.2, we present our empirical findings. In Section 4.3, we provide the cross-sectional results. In Section 4.4, we further explore the duration of overnight returns predictability for different stocks. Finally, Section 5 concludes our paper.

2 Literature Reviews

2.1 Theoretical Frameworks on Information of Non-trading Periods

Oldfield and Rogalski (1980) propose a model with multiple jump processes to explain the differences between returns in the trading and non-trading periods. They assume that returns over the trading and non-trading periods can be characterized by a stochastic process with diffusion and multiple jumps:

$$\frac{\mathrm{d}P}{P} = \mu \mathrm{d}t + \sigma \mathrm{d}\mathcal{W} + \sum_i z_i \mathrm{d}\theta_i$$

where $\mathcal{W}$ is the Wiener process, z represents random changes in share prices resulting from various non-trading periods' jumps, and $\mathrm{d}\theta$ is a random jump process associated with each jump. Hong and Wang (2000) provide an equilibrium model to describe a stock market with periodic market closures, which is able to explain the phenomenon of higher volatility during the trading period.

2.2 Empirical Findings on Information of Non-trading Periods

Fama (1965), French and Roll (1986), Lockwood and Linn (1990) show that the volatility in the trading period is much higher than the volatility in the nontrading period. French and Roll (1986) link this phenomenon to public and private information flow, and pricing errors. Barclay, Litzenberger, and Warner (1990) examine the Japanese stock market[1] to identify the determinants of stock return variance. The change in Japanese stock market trading days permits them to compare stock variances between trading Saturday and non-trading Saturday. They conclude that the key determinant of variance is private information. They find that neither the irrational trading noise hypothesis nor the public information hypothesis can explain the variance of the Japanese stock market return. Craig,

[1] Japanese stock market is allowed to trade for half a normal trading day approximately three Saturdays per month, and closed on other Saturdays.

Dravid, and Richardson (1995) use the Japanese Nikkei index-based futures traded in US as the proxy of contemporaneous overnight information in Japan stock market, and find close links between implied changes in the Nikkei index and actual overnight changes in the Nikkei index. They show that overnight information is efficiently reflected by derivatives traded in the US market. Barclay and Hendershott (2003) explore the relationship between after-hours trading and price discovery on Nasdaq. They find that albeit after-hour trading volume is low, it generates significant price discovery. They also find that information asymmetry in the pre-open period is higher than any other time of the day.

An important question is whether overnight return can predict daytime returns. Past studies have shown that overnight return has some predictability on daytime returns. Using cross listed firms in different exchanges, Chan, Chockalingam, and Lai (2000) show that overnight returns have significant effects on intra-day returns for the first half hour after market is open. By splitting close-to-close returns into two non-overlap series and introducing a functional coefficient model, Hong, Gallo, and Lee (2001) find evidence of an in-sample nonlinear predictability for daytime returns but this predictability is rejected by the out-sample test. In addition to return predictions, the issue of predictability of volatility has caught much attention since volatility is important for pricing derivatives. Gallo and Pacini (1998) show that the overnight surprise measured by absolute returns during market closures affects volatility. However, their finding is ambiguous because they used close-to-close returns instead of open-to-close returns to test the prediction of overnight shocks.

Our study is also related to the literature of public information effects. Ederington and Lee (1993) examine the impact of macroeconomic news releases on returns and volatility. They find that while returns reflect the information in less than a minute, the impact of news on volatility can persist for several hours. Berry and Howe (1994) measure the public information flow by the number of news released by Reuter's News Service per unit of time. They find a positive moderate relationship between public information flow and trading volume, but an insignificant relationship with price volatility. They also show that overnight information has significant impact on the opening volume which is consistent with Gerety and Mulherin (1992) findings of a positive relationship between opening volume and unexpected overnight volatility.

Understanding the effect of information in the non-trading period is important also for practitioners. As pointed out by Taylor (2005), a Value-at-Risk (VaR) model based on a conditional volatility model that incorporates overnight information produces more accurate results. Boes, Drost, and Werker (2006) model the overnight price process by a single jump while the daytime price process is following the standard affine model. They find that overnight jumps accounted for about $1/4$ of total jump risk.

While there are a number of studies on the relationship between overnight and daytime returns, there are relatively few studies on the predictive ability of overnight return on daytime volatility. This paper attempts to fill this gap. In addition to tests of overnight return information, we examine the relationship between overnight returns and daytime volatility for stocks with different frequencies.

3 Empirical Models

The close-to-close return for a security is

$$r_t = \ln C_t - \ln C_{t-1}$$

We can decompose returns into two distinct components, namely overnight returns and daytime returns,

$$r_t = \underbrace{\ln C_t - \ln O_t}_{\text{Daytime Returns}} + \underbrace{\ln O_t - \ln C_{t-1}}_{\text{Overnight Returns}}$$

An important question is whether overnight return contains information that affects the daytime market performance, a dimension of which is return volatility. The easiest way is test this hypothesis is to examine the relationship between two return series directly by estimating the following regression:

$$|y_t| = \theta_0 + \theta_1 |x_{t'}| + \theta_2 |y_{t-1}| + \nu_t$$

where y_t is the daytime return at time t, x_t is the overnight return at time t' preceded the market open at time t; More specifically,

$$r_1, r_2, ..., r_t, ... = \underbrace{x_{1'}, y_1}_{r_1}, \underbrace{x_{2'}, y_2}_{r_2}, ..., \underbrace{x_{t'}, y_t}_{r_t}, ...$$

However, using the absolute value of daytime returns to represent daytime volatility and the absolute value of overnight returns to represent the overnight shock may induce measurement errors. To conduct a more robust test, we employ the stochastic volatility model (SV henceforth) first proposed by Taylor (1982). Compared to the GARCH model, one of the main advantages of the SV model lies in the more realistic feature it embraces. Since volatility itself cannot be precisely predicted, adding another random shock in the volatility equation makes more sense in modeling the data. Assuming volatility itself follows a stochastic process is therefore more appealing than the GARCH model that does not allow for random disturbance. An important difference between SV or GARCH model and simple linear regression model is that the former treats volatility as a state variable, which cannot be observed. This specification allows us to avoid the error associated with volatility measured by absolute returns.

3.1 The Model

The SV model can be written as

$$\begin{cases} \epsilon_t = \exp(0.5h_t)\eta_t & \eta_t \sim \mathcal{N}(0,1) \\ h_t = \alpha_0 + \alpha_1 h_{t-1} + \omega_t & \omega \sim \mathcal{N}(0,\sigma_\omega^2) \end{cases}$$

This model was proposed by Taylor in 1982 but only becomes popular after mid 1990s. A main reason is that estimation of the model is much more complicated compared to the ARCH and GARCH models proposed by Engle (1982) and Bollerslev (1986) respectively. The likelihood function of the SV model is

$$l(\theta; y) = \int_{\mathbb{R}+} f(y, \mathbf{h}|\theta) \mathrm{d}\mathbf{h}$$

, which involves a multidimensional integral with no analytical form.

Due to the intractability of the likelihood function in the SV model, researchers often use an alternative method which focuses on sampling the density $f(\theta, \mathbf{h}|\mathbf{y})$ to overcome the estimation problem. In a seminal work, Jacquier, Polson, and Rossi (1994) use a Bayesian approach to estimate the SV model and draw the inference based on the posterior distribution.

$$f(\theta, h_0, ..., h_T | y_1, ..., y_T) \propto f(\theta) \prod_t f(h_t|h_{t-1}, \theta) \prod_t f(y_t|h_t, \theta)$$

In this paper, posterior distributions are obtained by using *Gibbs Sampling*, a *Markov Chain Monte Carlo* (MCMC) technique. The idea behind the MCMC method is to produce variate from a given multivariate density by repeatedly sampling a *Markov Chain* whose invariant distribution is the target density of interest. The *Forward Filtering Backward Sampling* algorithm developed by Carter and

Kohn (1994) and Fruhwirth-Schnatter (1994) is applied to estimate the time series of h_t based on draws of parameters in each step. Details of the estimation procedure are presented in the Appendix. In empirical estimation, we use 7 normal distributions with different means and variances to approximate the distribution of $\ln \chi_1^2$, which are also used by Kim, Shephard, and Chib (1998) to make the SV Estimation more robust. The 7 normal distributions are outlined in Table 1.

3.2 The Test Hypothesis

We can express the return process under stochastic volatility as

$$y_t = \rho_0 + \rho_1 y_{t-1} + \epsilon_t$$

where y_t represents daytime returns,

$$\begin{cases} \epsilon_t = \exp(0.5h_t)\eta_t & \eta_t \sim \mathcal{N}(0,1) \\ h_t = \alpha_0 + \alpha_1 h_{t-1} + \beta|x_{t'}| + \omega_t & \omega \sim \mathcal{N}(0,\sigma_\omega^2) \end{cases}$$

where $|x_{t'}|$ captures overnight shocks. We name this model $\mathcal{M}_1$.

Our null hypothesis is

$$\mathbb{H}_0 : \beta = 0$$

and the alternative hypothesis is

$$\mathbb{H}_1 : \beta \neq 0$$

The test hypothesis is that overnight returns contain no information to predict daytime return volatility. Under the null hypothesis, the model becomes

$$y_t = \rho_0 + \rho_1 y_{t-1} + \epsilon_t$$

$$\begin{cases} \epsilon_t = \exp(0.5h_t)\eta_t & \eta_t \sim \mathcal{N}(0,1) \\ h_t = \alpha_0 + \alpha_1 h_{t-1} + \omega_t & \omega \sim \mathcal{N}(0,\sigma_\omega^2) \end{cases}$$

where overnight shocks are not taken into consideration. We define this model as $\mathcal{M}_0$.

Bayesian analysis is used in making statistical inference. The hypothesis can be evaluated by Highest Posterior Density Intervals (HPDI, henceforth) and by the *Posterior Odds Ratio*.

The statistical inference using the HPDI is straightforward. For example, a 95% HPDI indicates that researchers are 95% confident that β lies within the HPDI. Although the inference is similar to the concept of confidence intervals in the standard statistic test, the interpretation is quite different. While $p - value$ is referred to as the confidence interval in the standard test, $p - value$ is referred to the posterior distribution of the parameter in the Bayesian inference.

Although it is easy to apply, the HPDI does not have a firm probability justification. An alternative is to adopt the concept of *Posterior Odds Ratio* which can be expressed as

$$\frac{\Pr(\mathcal{M}_0|y)}{\Pr(\mathcal{M}_1|y)} = \frac{\Pr(\mathcal{M}_0)}{\Pr(\mathcal{M}_1)} \times B$$

where B is the *Bayes Factor*

$$B = \frac{f(y|\mathcal{M}_0)}{f(y|\mathcal{M}_1)} = \frac{\int_{\Theta_0} f(y|\theta_0, \mathcal{M}_0) f(\theta_0|\mathcal{M}_0) \mathrm{d}\theta_0}{\int_{\Theta_1} f(y|\theta_1, \mathcal{M}_1) f(\theta_1|\mathcal{M}_1) \mathrm{d}\theta_1}$$

and $\frac{\Pr(\mathcal{M}_0)}{\Pr(\mathcal{M}_1)}$ is the *Prior Odds Ratio* which is set equal to 1 to avoid any prejudiced prior. The *Bayes factor* is estimated based on the method developed by Chib (1995):

$$\begin{aligned} \ln f(y|\mathcal{M}_0) - \ln f(y|\mathcal{M}_1) = & \ln f(y|\mathcal{M}_0, \theta_0^*) + \ln f(\theta_0^*|\mathcal{M}_0) - \ln f(\theta_0^*|\mathcal{M}_0, y) \\ & - [\ln f(y|\mathcal{M}_1, \theta_1^*) + \ln f(\theta_1^*|\mathcal{M}_1) - \ln f(\theta_1^*|\mathcal{M}_1, y)] \end{aligned}$$

for any values of θ_0^* and θ_1^*.

In our paper, we use the posterior mean as the point estimator for parameters and posterior densities $f(\theta_i|\mathcal{M}_i, y)$ are estimated at θ_i by applying the Gaussian kernel to the posterior distribution of parameters. The likelihood function $f(y|\mathcal{M}_i, \theta_i)$ is estimated by the filtering simulation. That is, $\mathcal{L}(\theta) = \sum_t \frac{\sum_1^M \ln f(y_t|h_t^j)}{M}$ where $h_t^j \sim \mathcal{N}(\hat{\alpha}_0 + \hat{\alpha}_1 h_{t-1}^j, \sigma_\omega^2)$.

The rationale of using the *Posterior Odds Ratio*is that the Bayesian analysis allows researchers to attach probability to competing models given a particular dataset. The *Posterior Odds Ratio* indicates which model is more likely to be the true data generating process. For example, if the *Posterior Odds Ratio* is less than unity, we would conclude that given the data, $\mathcal{M}_1$ is favored. Hence, the model selection is determined by judging on the probability of each model being the true DGP instead of resorting to some arbitrary criteria.

3.3 Priors

Since the SV model is estimated by the MCMC method under the Bayesian framework, we need to specify the prior distribution for each parameter. We use flat priors in our paper, meaning that parameter prior distributions are governed by a large dispersion. More specifically,

$$
\begin{aligned}
\alpha_0 &\sim \mathcal{N}(0, 10^8) \\
\alpha_1 &\sim \mathcal{N}(0, 10^8) \\
\beta &\sim \mathcal{N}(0, 10^4) \\
\sigma_\omega^2 &\sim \mathcal{IG}(\frac{2 \times 10^{-8}}{2}, \frac{2 \times 10^{-8}}{2})
\end{aligned}
$$

Flat priors permit the information contained in priors and its influenced to be reduced in the empirical analysis.

We adopt the Normal-Inverse Gamma prior in the empirical estimation which is somewhat different from previous studies. Despite this, using different priors should generate similar point estimation and statistical inferences since the likelihood function contains the same information from the data.

To illustrate that differences in priors shall not affect the posterior analysis, we adopt two priors to estimate the model from a simulated dataset and compare their performance. Table 2 reports the results based on the Normal-Inverse Gamma prior whereas Table 3 reports the results based on the prior suggested by Kim, Shephard, and Chib (1998). Results show that both priors lead to a very close posterior conclusion in terms of the point estimation and HPDI inferences. Therefore, the posterior analysis is not affected by the prior chosen. Thus, we use the Normal-Inverse Gamma prior in our empirical estimation.

4 Data and Empirical Results

4.1 Data

We employ both market index and individual firm data in empirical tests. For the market index, we use the daily data of Dow Jones Industrial Average index (DJIA) over the period from January 1997 to December 2005. The DJIA index is the most watched stock market index, which captures the performance of the overall market very well. In addition to the aggregate market data, we employ individual firm data to examine cross-sectional variations in volatility predictability. Specifically, we examine the information content of overnight returns for stocks with different trading frequency. Following previous studies [see Easley et al. (1996)], we use volume as a proxy. We sort all stocks traded on the NYSE into ten groups based on their average daily trading volumes over the period from January 1997 to December 2005. We then randomly choose 30 stocks in the first, fifth, and eighth deciles. In addition to the daily data, we also collected intraday data for these stocks from TAQ to examine the duration of overnight predictability for stocks of different frequency. Intraday analysis provides more detailed information to substantiate the empirical results for individual firms. The DJIA index data are obtained from Bloomberg and individual stock data are collected from CRSP and TAQ.

4.2 The Information Content of the Overnight Return

Table 4 reports summary statistics of the DJIA data. As shown, the variance of daytime returns is much larger than the variance of overnight returns. The result is consistent with the finding of French and Roll (1986) that the variance in the non-trading period is much less than that in the trading period. As shown in the last two panels of Figure 1, the time series of the daytime return is quite different from that of the overnight return, suggesting that they have different generating processes.

Table 5 and Figure 2 report the estimation result, MCMC iteration records and posterior distributions of parameters from the following linear regression model,

$$|y_t| = \theta_0 + \theta_1|x_{t'}| + \theta_2|y_{t-1}| + \nu_t$$

θ_1 represents the impact of overnight information to daytime volatility. This coefficient is significant since the 95% HPDI($[2.0419, 3.6290]$) does not contain 0. The left panels of Figure 2 show parameter values in the MCMC iternations. The estimation converges after 16000 iterations where we discard the first 1000 sweeps which form the burn-in period.

Table 6 and Figure 3 report the estimation result, MCMC iterations, and the posterior distributions of parameters from the stochastic volatility model $\mathcal{M}_0$ without the overnight information. On the other hand, Table 7 and Figure 4 report the estimation result, MCMC iterations, and the posterior distributions of parameters from the stochastic volatility model $\mathcal{M}_1$ with the overnight information. As shown in the right panels of Figure 3 and Figure 4 for the posterior distribution of parameters, both models show that past volatility can predict future volatility since the probability mass of parameter α_1 does not include 0. Table 6 shows that the 95% HPDI of α_1 is $[0.9717, 0.9933]$, whereas Table 7 shows that the 95% HPDI of α_1 is $[0.9511, 0.9952]$. Both indicate that past volatility has a significant predictive power for future volatility. The left panels of both figures show the MCMC iterations. The estimation of parameters converges after 16000 iterations where the first 1000 sweeps are discarded in the burn-in period.

The estimation results for the SV model that include the overnight information is consistent with those of the simple linear regression model. As shown in Table 7, under flat priors[2] the point estimate

[2]Flat priors are referred to prior distributions with large dispersions. More information of priors is illustrated in the Appendix, p. 27.

of β (the posterior mean) is 0.5201 indicating that the marginal effect of a overnight information shock to daytime volatility is about a half of its magnitude. This effect is significant since the 95% HPDI for β is $[0.0963, 1.0134]$ which does not include 0. The results are in favor of model $\mathcal{M}_1$, the stochastic volatility model including the overnight return as an explanatory variable. Results strongly suggest that the overnight return contains important information that affects daytime volatility. The Bayesian method provides the exact finite sample performance[3]. As such, Figure 4 illustrates exact distributions for the corresponding parameters. We observe that the probability mass of β being positive is dominant over the probability of β being negative. From the posterior distribution of β, we have $\Pr(\beta > 0|I_t) = 99.12\%$.

Given $M = 5000$ filtering conditional volatility, we have

$$\begin{aligned} B &= 0.018 \\ \Pr(\mathcal{M}_0 : \beta = 0|I) &= 1.77\% \\ \Pr(\mathcal{M}_1 : \beta \neq 0|I) &= 98.23\% \end{aligned}$$

The *Posterior Odds Ratio*, which is equal to 0.018, implies that model $\mathcal{M}_1$ is favored by the *Rule of Thumb* as suggested by Koop (2005). More specifically, there is only a 1.77% chance that model $\mathcal{M}_0$ is favored whereas there is a 98.23% chance that model $\mathcal{M}_1$ is favored.

Empirical evidence strongly indicates that $\beta \neq 0$ and the null hypothesis is soundly rejected. Thus, the overnight return contains important information that predicts daytime return volatility.

Our findings are consistent with the contention that daytime volatility reflects not only the new information at the daytime but also the overnight information. This is largely due to no trading in the overnight period and so the overnight information is impounded into the return next day. Thus, jumps in the overnight return could be used to predict daytime volatility. Our empirical evidence also

[3]The exact posterior distribution is based on the prior specification

suggests that daytime return variance is larger than the overnight return variance, since the former incorporates the information in the trading and nontrading periods while the latter may only reflect the overnight information partially.

We next compare the result of the stochastic volatility model with that of the GARCH model. Table 8 reports the estimation result for the GARCH(1,1) model with an exogenous variable in the volatility equation. More specifically, we estimate the following GARCH model:

$$y_t = \rho_0 + \rho_1 y_{t-1} + \epsilon_t$$

$$\begin{cases} \epsilon_t = \sqrt{h_t}\nu_t & \nu_t \sim \mathcal{N}(0,1) \\ h_t = \omega + \delta_1\epsilon_t^2 + \delta_2 h_{t-1} + \gamma|x_{t'}| \end{cases}$$

where $|x_{t'}|$ represents the overnight information and h_t is daytime volatility. The results for the GARCH model are similar to those for the SV model, confirming that daytime volatility reflects the overnight information. The γ coefficient is significantly different from 0 under the 95% confidence interval. Figure 5 plots the estimated volatility for both the SV model and the GARCH(1,1) model over time. Table 9 shows that the SV model performs better since the adjusted R^2 is higher than that of the GARCH(1,1) model. The result is consistent with the finding of Kim, Shephard, and Chib (1998) that the SV model outperforms the GARCH(1,1) model in predicting return volatility. Thus, in the remainder of our analysis, we focus on the SV model in volatility forecasting.

4.3 Trading Frequency and Overnight Information

Previous studies [see Easley et al. (1996)] have shown trades of less frequently traded stocks contain more information than those of more frequently traded stocks. We next examine whether the effect of the overnight return differs across stocks of different frequency. If trades of less frequently traded stocks contain more information, we would expect that the information accumulated overnight will be reflected more in the daytime trading of less frequently traded stocks. We estimate the SV model for stocks in the high, medium and low trading frequency categories as described in Section 4.1.

The results of estimation for the three groups are reported in (Table 10, Table 11,and Table 12), respectively. These results conform our finding for the market index that the overnight return contain useful information to predict daytime return volatility. More importantly, we find that the predicting power of the overnight information is much higher for less frequently traded stocks. The high trading frequency group has the lowest marginal effect of the overnight information $\beta_i = 5\%$ while the low trading frequency group have the largest marginal effect of the overnight information, $\beta_i = 13.96\%$. Thus, the predictability of the overnight information on daytime return volatility is larger for less actively traded stocks.

The mid column of Figure 6 shows that more active stocks tend to have more overnight information than less active stocks as reflected by the absolute overnight returns. However, the marginal effect of the overnight information is the smallest for the most active stocks.

The lower marginal effect of the overnight information for more actively traded stocks is consistent with our prediction. On the one hand, market microstructure theory Easley et al. (1996) suggests that the probability of information-based trading is lower for more active stocks. Since the overnight information primarily contains the private information [see Barclay, Litzenberger, and Warner (1990), Lockwood and Linn (1990)], and informed trading intensity tends to be higher for less active stocks, Hasbrouck (1991), the overnight information should have higher predictability on the daytime return volatility of inactively traded stocks. On the other hand, more new information arrives during the daytime for active stocks than for inactive stocks. Also, since the daytime information is more than the overnight information, the former will eventually dominate the latter in determining daytime volatility. As shown in Figure 6, the magnitude for trading volume for the most actively traded group is dramatically larger than that for the least actively traded group. Therefore the daytime information should carry a heavier weight in determining daytime volatility in active stocks than in inactive stocks. Thus, the magnitude of the overnight information effect on the daytime volatility would tend to be smaller for active stocks.

Daytime volatility reflect both the daytime information and the overnight information. Since for high trading frequency group, more daytime information is impounded into those stocks, the predictability of the overnight information tends to be lower, even though its overnight jumps are large.

In summary, the cross-sectional study supports our conjecture that the overnight information has predictive power on daytime volatility. Furthermore, our results suggest that the overnight information plays a more influential role in inactive stocks than in active stocks in the prediction of daytime volatility.

4.4 Intraday Estimated Results

To provide a further diagnosis on the relationship between daytime volatility and the daytime and overnight information, we estimate the following regression model using the intraday trading data:

$$|y_{t_i}| = \alpha + \beta \times |x_{t'}| + \gamma \times |y_{t-1}| + \epsilon_{t_i}$$

, where

- $x_{t'}$ represents the overnight return at time t' preceded to the market open at time t
- y_{t-1} is the daytime return at time $t-1$
- y_{t_i} is the intraday return at time t, for example, time $t_{9:30to10:00A.M.}$ represents the first half hour return after market opens.

- t_i indexes the time of intraday transactions

Table 13 and Table 14 report the results of estimation. The estimate of β, which represents the marginal effect of overnight information, shows that the overnight information effect on both groups decreases over the daytime trading period with the highest on the first half hour of trading and the lowest near the market closure. Figure 7 shows the evolution of β, which depicts a gradual assimilation process on overnight information. There are differences between actively and inactively traded stocks. For actively traded stocks, the decline in the marginal effect of the overnight information is greater right after the market open.

Table 15 and Table 16 summarize the t-statistics of β for both groups. In Table 15, the median t-statistic is significant up until 11:00 a.m. for active traded stocks, whereas Table 16 shows that the median t-statistic remains significant in the afternoon. Results suggest that the effect of the overnight information on return volatility is more persistent for inactively traded stocks.

The higher diminishing rate for the marginal effect β of the overnight information for active traded stocks suggests that the overnight information is impounded into prices more quickly for these stocks due to more active trading. It also reflects that daytime information is assimilated faster into actively traded stocks. Previous studies have shown that the information arrival rate is higher during the daytime trading period. High trading frequency tends to reveal the information faster for actively traded stocks, thereby further dilute the effect of the overnight information.

Results show that the overnight information plays a more important role in affecting daytime volatility for inactive stocks. Because these stocks are less frequently traded, the effect of the overnight information lasts longer. At the same time, the impounding of the daytime information is slower. Both forces tend to increase the relative effect of the overnight information. As a consequence, the predictive power of the overnight information for daytime return volatility tends to be higher for inactive stocks.

5 Conclusion

In this paper, we examine the predictability of the overnight information on daytime return volatility. Information is accumulated overnight, but there is no trading to reflect the information. The accumulated information thus tends to reflected in the daytime return volatility. Both of our linear regression model and the SV model suggest that overnight shocks, measured as absolute value of overnight returns, have significant impact on prediction of daytime volatility.

In addition to the test for the information effect at the market level, we examine individual stocks with different trading frequencies. Market microstructure theory suggests that higher trading volume stocks contain less informed trading than lower trading volume stocks. Since the overnight information mainly consists of private information[4], trading of the inactive stocks during the daytime contains a higher private information content that affects return volatility. Empirical evidence shows that the marginal effect of overnight information shocks in inactive stocks is larger. This result is consistent with previous findings that there is more intense arrivals of new information during the daytime and slower release of the overnight information by inactive stocks. Empirical findings based on intraday transactions confirm this argument and suggest that the overnight information plays a more important role in affecting the daytime volatility of inactively traded stocks. These in turn suggest that for inactive stocks, the overnight information has more predictive power on return volatility. The effect of the overnight information lasts longer for inactive stocks than in active stocks, implying that more frequently traded stocks impound the information more efficiently.

[4] Barclay, Litzenberger, and Warner (1990), Lockwood and Linn (1990) have explicitly test the private Information-based model and conclude that non-trading information generally consists of private information.

References

Barclay, Michael J., and Terrence Hendershott. 2003. "Price discovery and trading after hours." *Review of Financial Studies* 16:1041–1073.

Barclay, Michael J., Robert H. Litzenberger, and Jerold B. Warner. 1990. "Private information, trading volume, and stock-return variances." *Review of Financial Studies* 3:233–253.

Berry, Thomas D., and Keith M. Howe. 1994. "Public information arrival." *Journal of Finance* 49:1331–1346.

Boes, Mark-Jan, Feike C. Drost, and Bas J.M. Werker. 2006. "The impact of overnight periods on option pricing." working paper.

Bollerslev, Tim. 1986. "Generalized autoregressive conditional heteroskedasticity." *Journal of Econometrics* 31:307–327.

Carter, Chris K., and Robert Kohn. 1994. "On Gibbs sampling for state space models." *Biometrika* 81:541–533.

Chan, Kalok, Mark Chockalingam, and Kent W.L. Lai. 2000. "Overnight information and intraday trading behavior: evidence from NYSE cross-listed stocks and their local market information." *Journal of Multinational Financial Management* 10:495–509.

Chib, Siddhartha. 1995. "Marginal likelihood from the Gibbs output." *Journal of the American Statistical Association* 90:1313–1321.

Craig, Alastair, Ajay Dravid, and Matthew Richardson. 1995. "Market efficiency around the clock some supporting evidence using foreign-based derivatives." *Journal of Financial Economics* 39:161–180.

Dufour, Alfonso, and Robert F. Engle. 2000. “Time and the price impact of a trade.” *Journal of Finance* 55:2467–2498.

Easley, David, Nicholas M.Kiefer, Maureen O’Hara, and Joseph B.Paperman. 1996. “Liquidity, information, and infrequently traded stocks.” *The Journal of Finance* 51:1405–1436.

Ederington, Louis H., and Jae Ha Lee. 1993. “How markets process information: news releases and volatility.” *Journal of Finance* 48:1161–1191.

Engle, Robert. 1982. “Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation.” *Econometrica* 50:987–1008.

Fama, Eugene F. 1965. “The behavior of stock-market prices.” *Journal of Business* 38:34–105.

French, Kenneth R., and Richard Roll. 1986. “Stock return variances the arrival of information and the reaction of traders.” *Journal of Financial Economics* 17:5–26.

Fruhwirth-Schnatter. 1994. “Data augmentation and dynamic linear models.” *Journal of Time Series Analysis* 15:183–202.

Gallo, Giampiero M., and Barbara Pacini. 1998. “Early news is good news: the effects of market opening on market volatility.” *Studies in Nonlinear Dynamics and Economics* 2:115–131.

Gerety, Mason S., and J. Harold Mulherin. 1992. “Trading halts and market activity: an analysis of volume at the open and the close.” *Journal of Finance* 47:1765–1784.

Granger, Clive W.J., and Ser-Huang Poon. 2003. “Forecasting volatility in financial markets: a review.” *Journal of Economic Literature* 41:478–539.

Hasbrouck, Joel. 1991. “Measuring the information content of stock trades.” *Journal of Finance* 46:179–207.

Hong, Harrison, and Jiang Wang. 2000. "Trading and returns under periodic market closures." *Journal of Finance* 55:297–354.

Hong, Yongmiao, Giampiero M. Gallo, and Tae-Hwy Lee. 2001. "Modelling the impact of overnight surprises on intra-daily stock returns." working paper.

Jacquier, Eric, Nicholas G. Polson, and Peter E. Rossi. 1994. "Bayesian analysis of stochastic volatility models." *Journal of Business and Economic Statistics* 12:371–389.

Kim, Sangjoon, Neil Shephard, and Siddhartha Chib. 1998. "Stochastic volatility: likelihood inference and comparison with ARCH models." *Review of Economics Studies* 65:361–393.

Koop, Gary. 2005. *Bayesian econometrics*. Edited by Gary Koop. Wiley.

Lockwood, Larry J., and Scott C. Linn. 1990. "An examination of stock market return volatility during overnight and intraday periods, 1964-1989." *Journal of Finance* 45:591–601.

Oldfield, George S., and Richard J. Rogalski. 1980. "A theory of common stock returns over trading and non-trading periods." *Journal of Finance* 35:729–751.

Taylor, Nicholas. 2005. "The importance of overnight information in risk management models." working paper.

Taylor, Stephen J. 1982. *Time series analysis: theory and practise*. Volume 1. Edited by O.D.Anderson. Elsevier.

West, Mike, and Jeff Harrison. 1997. *Bayesian forecasting and dynamic models*. Springer-Verlag New York, Inc.

Xu, Xiaoqing Eleanor, Peter Chen, and Chunchi Wu. 2006. "Time and dynamic volume-volatility relation." *Journal of Banking and Finance* 30:1535–1558.

Appendix

Univariate Dynamic Linear Model with Exogenous Variables

For Univariate Dynamic Linear Model,

$$\begin{cases} y_t = x_t'\beta + F'\xi_t + \epsilon_t & \epsilon_t \sim \mathcal{N}(0, \sigma_\epsilon^2) \\ \xi_t = G\xi_{t-1} + \psi z_t + \omega_t & \omega_t \sim \mathcal{N}(0, \sigma_\omega^2) \end{cases}$$

Let I_t represent the information up to time t,in other words $I_t = y_0, ..., y_t; x_0, ..., x_t; z_0, ..., z_t, z_{t+1}$ by linear projection we would have following results [see West and Harrison (1997)]

- the *forward filtering*,

  1. Initialization

$$\begin{aligned} \hat{\xi}_{1|0} &= E(\xi_1) \\ P_{1|0} &= E\left[\xi_1 - E(\xi_1)\right]^2 \end{aligned}$$

  2. the *Evolution* step

$$\begin{aligned} \hat{\xi}_{t|t-1} &= G\xi_{t-1|t-1} + \psi z_t \\ P_{t|t-1} &= GP_{t-1|t-1}G' + \sigma_\omega^2 \end{aligned}$$

  3. the *Updating* step

$$\begin{aligned} \hat{\xi}_{t|t} &= \hat{\xi}_{t|t-1} + P_{t|t-1}F(F'P_{t|t-1}F + \sigma_\epsilon^2)^{-1}(y_t - x_t'\beta - F'\xi_t) \\ P_{t|t} &= P_{t|t-1} - P_{t|t-1}F(F'P_{t|t-1}F + \sigma_\epsilon^2)^{-1}F'P_{t|t-1} \end{aligned}$$

where

$$\begin{aligned}
\hat{\xi}_{t|t-1} &= E(\xi_t|I_{t-1}) \\
\hat{\xi}_{t|t} &= E(\xi_t|y_t, x_t, I_{t-1}) \\
P_{t|t} &= E[(\xi_t - \hat{\xi}_{t|t})(\xi_t - \hat{\xi}_{t|t})'] \\
P_{t|t-1} &= E[(\xi_t - \hat{\xi}_{t|t-1})(\xi_t - \hat{\xi}_{t|t-1})']
\end{aligned}$$

Results above are named *forward filtering*, and for *backward sampling*.

- the *backward* distribution is

$$(\xi_{t-1}|\xi_t, I_t) \sim \mathcal{N}(m_t, M_t)$$

where

$$\begin{aligned}
B_{t-1} &= P_{t-1|t-1} G' P_{t|t-1}^{-1} \\
m_t &= \xi_{t|t} + B_{t-1} \times (\xi_t - \hat{\xi}_{t|t-1}) \\
M_t &= P_{t-1|t-1} - B_{t-1} P_{t|t-1} B_{t-1}'
\end{aligned}$$

Proof. Using Bayes' theorem, we could have

$$f(\xi_{t-1}|\xi_t, I_{t-1}) \propto f(\xi_{t-1}|I_{t-1}) f(\xi_t|\xi_{t-1}, I_{t-1})$$

and since

$$\begin{aligned}
(\xi_{t-1}|I_{t-1}) &\sim \mathcal{N}(\hat{\xi}_{t-1|t-1}, P_{t-1|t-1}) \\
(\xi_t|\xi_{t-1}, I_{t-1}) &\sim \mathcal{N}(\hat{\xi}_{t|t-1}, P_{t|t-1})
\end{aligned}$$

by simple algebra we would have

$$
\begin{aligned}
(\xi_{t-1}|\xi_t, I_{t-1}) &\sim \mathcal{N}(m_t, M_t) \\
B_{t-1} &= P_{t-1|t-1} G' P_{t|t-1}^{-1} \\
m_t &= \xi_{t|t} + B_{t-1} \times (\xi_t - \hat{\xi}_{t|t-1}) \\
M_t &= P_{t-1|t-1} - B_{t-1} P_{t|t-1} B_{t-1}'
\end{aligned}
$$

■

SV models could be transformed into DLM, which would be shown in followings.

$$
\begin{cases}
\epsilon_t = \exp(0.5 h_t)\eta_t & \eta_t \sim \mathcal{N}(0,1) \\
h_t = \alpha_0 + \alpha_1 h_{t-1} + \beta X_t + \omega_t & \omega_t \sim \mathcal{N}(0, \sigma_\omega^2)
\end{cases}
$$

$\Rightarrow$

$$
\begin{cases}
\ln \epsilon_t^2 = h_t + \ln \eta_t^2 & \ln \eta_t^2 \sim \ln \chi_1^2 \\
h_t = \alpha_0 + \alpha_1 h_{t-1} + \beta X_t + \omega_t & \omega_t \sim \mathcal{N}(0, \sigma_\omega^2)
\end{cases}
$$

and the $\ln \chi_1^2$ is approximated by 7 normal distributions

$$
\ln \chi_1^2 \overset{a}{\sim} \sum_1^7 w_i \mathcal{N}(\mu_i, \tau_i^2)
$$

which can be found in Table 1.

An Artificial Data Generating Process

A simulated example of the canonical stochastic volatility model with external variables in state equations.

$$
\begin{cases}
y_t = \exp(0.5 h_t)\sigma_t & \sigma_t \sim \mathcal{N}(0,1) \\
h_t = 0.95 \times h_{t-1} + 1.5 \times X_t + \nu_t & \nu_t \sim \mathcal{N}(0, 0.15^2)
\end{cases}
$$

where $\alpha_0 = 0$,$\alpha_1 = 0.95$,$\sigma_\omega = 0.15$,and $\beta = 1.5$. Results are shown in Table 2.

Adopting alternative prior beliefs

Instead of using Normal-InverseGamma priors we also adopt priors used in the Kim, Shephard, and Chib (1998).

$$\begin{cases} \epsilon_t = \exp(0.5h_t)\mu_t & \mu_t \sim \mathcal{N}(0,1) \\ h_t = \mu + \phi(h_{t-1} - \mu) + \psi X_t + \eta & \eta \sim \mathcal{N}(0, \sigma_\eta^2) \end{cases}$$

and priors are,

$$\begin{aligned} \mu &\sim \mathcal{N}(0,10) \\ \phi &= 2\phi^* - 1 \\ \phi^* &\sim Beta(20, 1.5) \\ \sigma_\eta &\sim \mathcal{IG}(2.5, 0.025) \\ \psi &\sim \mathcal{N}(0,10) \end{aligned}$$

The estimated results for the artificially generated data based on 16000 sweeps where first 5000 acts as the burn-in period. Results are included in Table 3.

Table 1: Using 7 normal distribution to approximate $\ln \chi_1^2$

i	w	μ	τ^2
1	0.00730	−11.40039	5.79596
2	0.10556	−5.24321	2.61369
3	0.00002	−9.83726	5.17950
4	0.04395	1.50746	0.16735
5	0.34001	−0.65098	0.64009
6	0.24566	0.52478	0.34023
7	0.25750	−2.35859	1.26261

Table 2: Results of the Simulated Data

	Mean	Std.	2.5%	median	97.5%
α_0	−0.0023	0.0044	−0.0112	−0.0022	0.0064
α_1	0.9519	0.0042	0.9435	0.9520	0.9598
β	1.5413	0.0889	1.3670	1.5412	1.7255
σ_w	0.1176	0.0225	0.0772	0.1158	0.1664

$$\begin{cases} y_t = \exp(h_t/2)\sigma_t & \sigma_t \sim \mathcal{N}(0,1) \\ h_t = \alpha_0 + \alpha_1 h_{t-1} + \beta X_{t'} + \nu_t & \nu_t \sim \mathcal{N}(0,\sigma_\nu^2) \end{cases}$$

Table 3: Results Based on Alternative Priors

	Mean	Std.	2.5%	median	97.5%
μ	−0.059	0.0856	−0.2292	−0.0576	0.1068
ϕ	0.9528	0.0043	0.9441	0.9529	0.9608
ψ	1.532	0.0923	1.387	1.519	1.759
σ_η	0.1201	0.0157	0.0942	0.1178	0.1555

$$\begin{cases} \epsilon_t = \exp(h_t/2)\mu_t & \mu_t \sim \mathcal{N}(0,1) \\ h_t = \mu + \phi(h_{t-1} - \mu) + \psi X_{t'} + \eta_t & \eta_t \sim \mathcal{N}(0, \sigma_\eta^2) \end{cases}$$

Table 4: Summary Statistics of DJIA data

	Observations	Mean	Variance
Close-to-Close Returns	2264	0.0225	1.3416
Open-to-Close Returns	2264	0.0298	1.3217
Close-to-Open Returns	2264	−0.0073	0.0022

Table 5: OLS Estimation of DJIA

	Mean	Std.	2.5%	median	97.5%
θ_0	0.6696	0.0255	0.6194	0.6695	0.7197
θ_1	2.8431	0.4020	2.0419	2.8440	3.6290
θ_2	0.1166	0.0210	0.0755	0.1167	0.1572
θ_3	0.7691	0.0114	0.7472	0.7690	0.7919

$$|y_t| = \theta_0 + \theta_1 |x_{t'}| + \theta_2 |y_{t-1}| + \nu_t$$

where $x_{t'}$ represents the overnight return on time t', and y_t represents the daytime return on time t.

Table 6: Results of Stochastic Volatility Model without Overnight Shocks

	Mean	Std.	2.5%	median	97.5%
α_0	-0.0007	0.0028	-0.0063	-0.0006	0.0049
α_1	0.9835	0.0056	0.9717	0.9840	0.9933
σ_w	0.1314	0.0170	0.1017	0.1308	0.1663

$$y_t = \rho_0 + \rho_1 y_{t-1} + \epsilon_t$$

$$\begin{cases} \epsilon_t = \exp(h_t/2)\eta_t & \eta_t \sim \mathcal{N}(0,1) \\ h_t = \alpha_0 + \alpha_1 h_{t-1} + \omega_t & \omega_t \sim \mathcal{N}(0, \sigma_\omega^2) \end{cases}$$

where y_t is the daytime return and h_t is the conditional daytime volatility.

Table 7: Results of Stochastic Volatility Model with Overnight Shocks

	Mean	Std.	2.5%	median	97.5%
α_0	-0.0156	0.0434	-0.1025	-0.0151	0.0688
α_1	0.9752	0.0111	0.9511	0.9759	0.9952
β	0.5201	0.2324	0.0963	0.5088	1.0134
σ_w	0.1411	0.0209	0.1050	0.1392	0.1882

$$y_t = \rho_0 + \rho_1 y_{t-1} + \epsilon_t$$

$$\begin{cases} \epsilon_t = \exp(h_t/2)\eta_t & eta_t \sim \mathcal{N}(0,1) \\ h_t = \alpha_0 + \alpha_1 h_{t-1} + \beta|x_{t'}| + \omega_t & \omega_t \sim \mathcal{N}(0, \sigma_\omega^2) \end{cases}$$

where y_t and h_t is same as before, $x_{t'}$ is the overnight return.

Table 8: Results of GARCH(1,1) with Overnight Shocks

	Coef	Std.	t-Stat	Pr
ω	0.0098	0.0050	1.9391	0.0525
δ_1	0.0877	0.0068	12.9800	0.0000
δ_2	0.8988	0.0094	95.7356	0.0000
γ	0.4690	0.1946	2.4104	0.0159

$$y_t = \rho_0 + \rho_1 y_{t-1} + \epsilon_t$$

$$\begin{cases} \epsilon_t = \sqrt{h_t}\nu_t & \nu_t \sim \mathcal{N}(0,1) \\ h_t = \omega + \delta_1 \epsilon_t^2 + \delta_2 h_{t-1} + \gamma |x_{t'}| \end{cases}$$

where y_t is the daytime return and $x_{t'}$ is the overnight return and h_t is the conditional daytime volatility.

Table 9: Predictability Comparisons with GARCH using $|y_t|$

	Const	$E(h_t\|I_t)$	$100 \times Adj.R^2$
SV Model	-0.1953^{***}	0.9113^{***}	35.39
GARCH(1,1) Model	0.5344^{***}	0.2251^{***}	10.46

$$|y_t| = Const + Coef \times E(h_t|I_t) + e_t$$

where y_t is the daytime return and $E(h_t|I_t)$ is the estimated daytime volatility from SV and GARCH model respectively.

Table 10: Results for High Trading Frequency Sample

	Ticker	Mean	Std.	2.5%	median	97.5%
1	A	0.0172	0.0068	0.0046	0.0170	0.0311
2	ADI	0.0205	0.0064	0.0088	0.0203	0.0339
3	ADM	0.2747	0.0402	0.2002	0.2729	0.3571
4	AMD	0.1068	0.0165	0.0772	0.1056	0.1438
5	AMP	0.1022	0.0612	0.0034	0.0952	0.2321
6	ANDW	0.0766	0.0189	0.0431	0.0752	0.1180
7	ATI	0.0016	0.0003	0.0010	0.0016	0.0024
8	AV	0.0178	0.0050	0.0087	0.0176	0.0281
9	CAT	0.0160	0.0058	0.0049	0.0159	0.0279
10	CCU	0.0332	0.0093	0.0161	0.0330	0.0524
11	CHIR	0.0314	0.0109	0.0110	0.0312	0.0529
12	CNC	−0.0056	0.0025	−0.0102	−0.0058	−0.0002
13	CVS	0.0202	0.0065	0.0077	0.0199	0.0336
14	DISH	0.0290	0.0061	0.0175	0.0290	0.0412
15	DOW	0.0255	0.0081	0.0112	0.0249	0.0433
16	DUK	0.0167	0.0098	−0.0009	0.0161	0.0371
17	ERICY	0.0013	0.0025	−0.0029	0.0011	0.0066
18	FISV	0.0298	0.0104	0.0094	0.0296	0.0504
19	FRE	0.0051	0.0048	−0.0044	0.0051	0.0145
20	HLT	0.0884	0.0230	0.0486	0.0869	0.1382
21	HOT	0.0530	0.0177	0.0235	0.0512	0.0928
22	NOK	0.0193	0.0043	0.0113	0.0191	0.0282
23	SPY	0.2132	0.0382	0.1416	0.2111	0.2919
24	TJX	0.0135	0.0045	0.0048	0.0135	0.0228
25	TXU	0.0425	0.0145	0.0173	0.0413	0.0753
26	VIA	0.0811	0.0154	0.0527	0.0803	0.1133
27	VOD	0.0140	0.0035	0.0077	0.0137	0.0218
28	WMB	0.0577	0.0124	0.0358	0.0568	0.0842
29	X	0.0985	0.0309	0.0400	0.0963	0.1636
30	YHOO	0.0017	0.0037	−0.0052	0.0015	0.0095

Table 11: Results for Mid Trading Frequency Sample

	Ticker	Mean	Std.	2.5%	median	97.5%
1	ABN	−0.0090	0.0010	−0.0110	−0.0089	−0.0069
2	AIRT	0.1227	0.0144	0.0949	0.1226	0.1514
3	ALAN	0.0686	0.0106	0.0490	0.0690	0.0877
4	ASHW	0.1240	0.0254	0.0767	0.1231	0.1748
5	ATLS	0.1678	0.0698	0.0559	0.1540	0.3181
6	BBG	0.0658	0.0440	−0.0158	0.0648	0.1573
7	BLK	0.0465	0.0154	0.0173	0.0461	0.0782
8	BPT	0.1563	0.0259	0.1072	0.1555	0.2100
9	CMTL	0.0922	0.0127	0.0672	0.0920	0.1170
10	CNLG	0.0431	0.0065	0.0309	0.0430	0.0561
11	CPTS	0.0885	0.0114	0.0678	0.0881	0.1119
12	CTRX	0.0228	0.0081	0.0091	0.0222	0.0407
13	DUSA	0.1387	0.0189	0.1037	0.1384	0.1770
14	ENZ	0.0926	0.0224	0.0538	0.0906	0.1411
15	EWS	0.1021	0.0184	0.0687	0.1009	0.1409
16	EWW	0.1649	0.0272	0.1158	0.1637	0.2207
17	GFF	0.1686	0.0333	0.1055	0.1679	0.2366
18	HCN	0.2871	0.0493	0.1955	0.2856	0.3879
19	HH	0.0623	0.0133	0.0380	0.0619	0.0901
20	ISCA	0.1927	0.0369	0.1258	0.1911	0.2702
21	JOSB	0.1056	0.0173	0.0726	0.1053	0.1403
22	PGI	0.0050	0.0035	−0.0014	0.0048	0.0121
23	SUG	0.0839	0.0198	0.0488	0.0826	0.1282
24	TXI	0.0610	0.0250	0.0195	0.0584	0.1187
25	USA	0.1838	0.0296	0.1286	0.1824	0.2456
26	VMSI	0.0346	0.0108	0.0141	0.0342	0.0564
27	VSAT	0.0488	0.0130	0.0257	0.0480	0.0757
28	WRI	0.0825	0.0202	0.0461	0.0817	0.1257
29	WSTM	0.1111	0.0178	0.0729	0.1129	0.1415
30	ZONA	0.0490	0.0100	0.0309	0.0484	0.0699

Table 12: Results for Low Trading Frequency Sample

	Ticker	Mean	Std.	2.5%	median	97.5%
1	AAON	0.1986	0.0294	0.1431	0.1976	0.2564
2	APX	0.2753	0.0518	0.1804	0.2723	0.3839
3	ASGR	0.1248	0.0220	0.0825	0.1254	0.1659
4	BNS	0.0040	0.0042	−0.0039	0.0038	0.0128
5	BPL	0.0666	0.0255	0.0196	0.0652	0.1187
6	BTF	0.1873	0.0352	0.1215	0.1865	0.2591
7	CEE	0.2103	0.0426	0.1406	0.2046	0.3088
8	CFS	0.1421	0.0121	0.1189	0.1420	0.1665
9	DASTY	0.0306	0.0074	0.0173	0.0301	0.0462
10	DDRX	0.0764	0.0088	0.0591	0.0763	0.0940
11	DRCO	0.1401	0.0183	0.1051	0.1398	0.1770
12	EMBX	0.1954	0.0224	0.1533	0.1948	0.2408
13	FFEX	0.1092	0.0172	0.0770	0.1085	0.1441
14	FTF	0.0649	0.0671	−0.0588	0.0626	0.2029
15	GBCI	0.0927	0.0218	0.0528	0.0917	0.1379
16	GBX	0.2857	0.0477	0.1928	0.2850	0.3795
17	GSH	0.1213	0.0153	0.0934	0.1206	0.1535
18	IF	0.1071	0.0226	0.0678	0.1053	0.1500
19	IMKTA	0.2131	0.0314	0.1534	0.2133	0.2750
20	LANV	0.0830	0.0089	0.0655	0.0828	0.1007
21	LBF	0.2248	0.0332	0.1626	0.2233	0.2944
22	PHF	0.2534	0.0492	0.1625	0.2514	0.3549
23	SIMC	0.1111	0.0136	0.0856	0.1106	0.1390
24	TKF	0.1161	0.0207	0.0778	0.1150	0.1600
25	TRH	0.0379	0.0237	−0.0058	0.0372	0.0869
26	UMBF	0.3324	0.0492	0.2373	0.3324	0.4279
27	USAK	0.1534	0.0199	0.1151	0.1529	0.1941
28	WRS	0.0160	0.0366	−0.0481	0.0133	0.0982
29	WST	0.1348	0.0356	0.0759	0.1291	0.2126
30	ZIF	0.0790	0.0239	0.0367	0.0777	0.1303

Table 13: Summary of β for High Trading Frequency Group

	09:30-10:00	10:00-11:00	11:00-12:00	12:00-13:00	13:00-14:00	14:00-15:00
average β	0.0349	0.0162	0.0126	0.0080	0.0072	0.0080
median β	0.0182	0.0085	0.0049	0.0038	0.0034	0.0036
25% β	0.0044	0.0044	0.0022	0.0006	0.0016	0.0007
95% β	0.0464	0.0156	0.0122	0.0076	0.0053	0.0081

Table 14: Summary of β for Low Trading Frequency Group

	09:30-10:00	10:00-11:00	11:00-12:00	12:00-13:00	13:00-14:00	14:00-15:00
average β	0.0241	0.0139	0.0078	0.0052	0.0058	0.0057
median β	0.0196	0.0069	0.0043	0.0035	0.0033	0.0037
25% β	0.0064	0.0032	0.0016	0.0016	0.0009	0.0012
95% β	0.0292	0.0144	0.0084	0.0064	0.0078	0.0082

Table 15: Summary of t Statistics for High Trading Frequency Group

	09:30-10:00	10:00-11:00	11:00-12:00	12:00-13:00	13:00-14:00	14:00-15:00
average t Stat.	4.8979	3.4401	2.9310	2.1504	2.0731	1.8953
median t Stat.	2.8844	2.2478	1.8814	1.4546	1.6222	1.5533
25% t Stat.	1.7438	1.0916	0.6881	0.4831	0.9350	0.6335
95% t Stat.	6.0208	4.7415	3.3743	2.7044	2.2114	2.9616

Table 16: Summary of t Statistics for Low Trading Frequency Group

	09:30-10:00	10:00-11:00	11:00-12:00	12:00-13:00	13:00-14:00	14:00-15:00
average t Stat.	8.0775	5.2193	3.2167	2.6105	2.5823	2.7694
median t Stat.	7.6459	5.0384	2.8017	1.8841	2.7932	2.9824
25% t Stat.	3.5928	2.0193	1.3332	1.0591	0.9299	0.9672
95% t Stat.	12.3445	6.9451	5.1569	4.2514	3.8299	4.0827

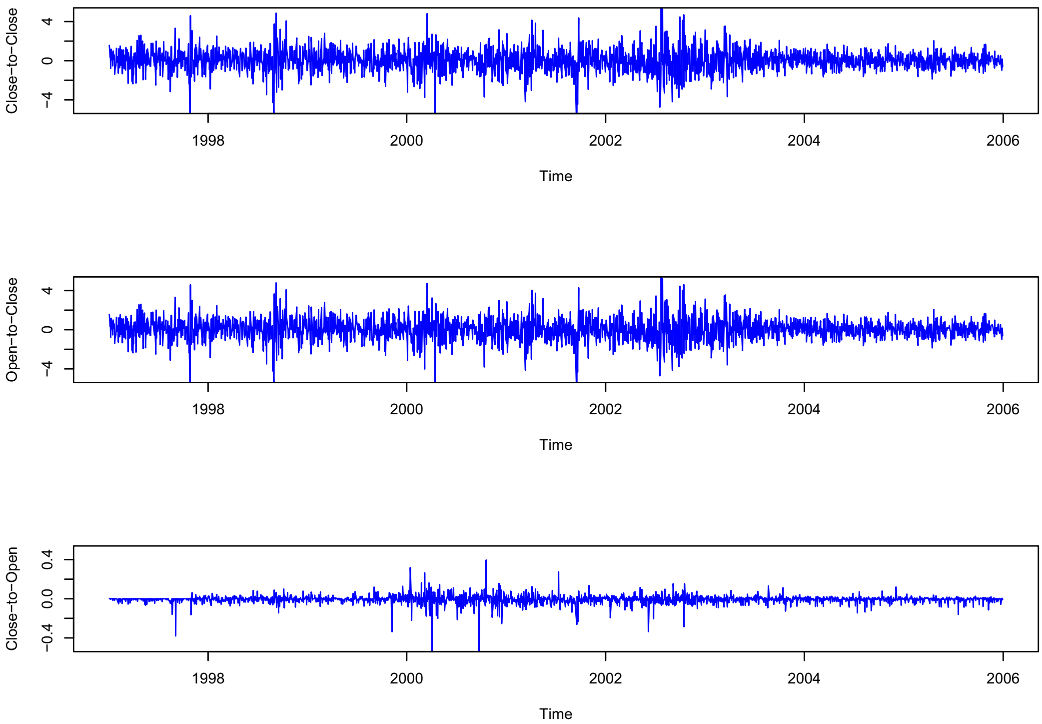

Figure 1: Time Series for Trading and Non-trading Periods

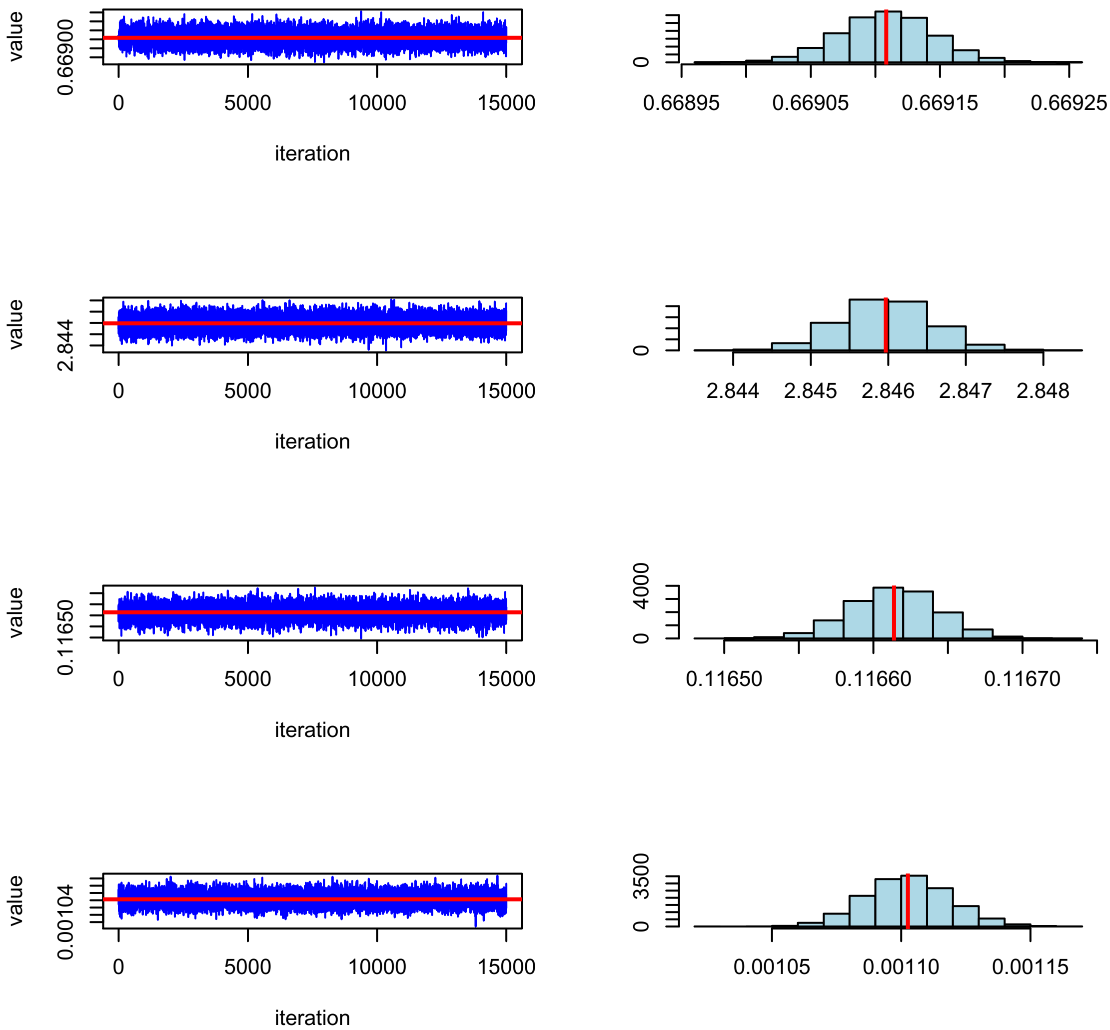

Figure 2: Results of the OLS Model

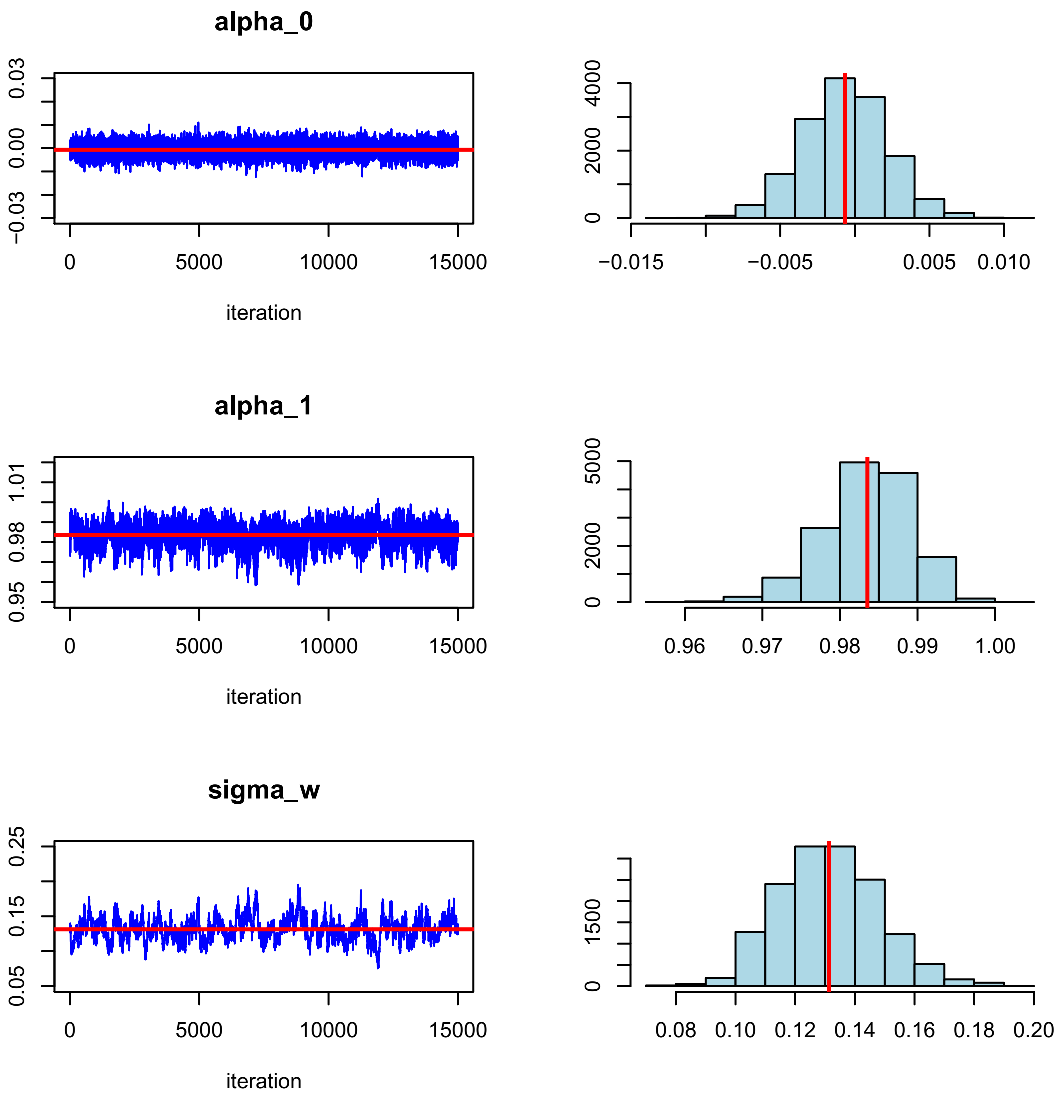

Figure 3: Results of Stochastic Volatility Model without Overnight Shocks

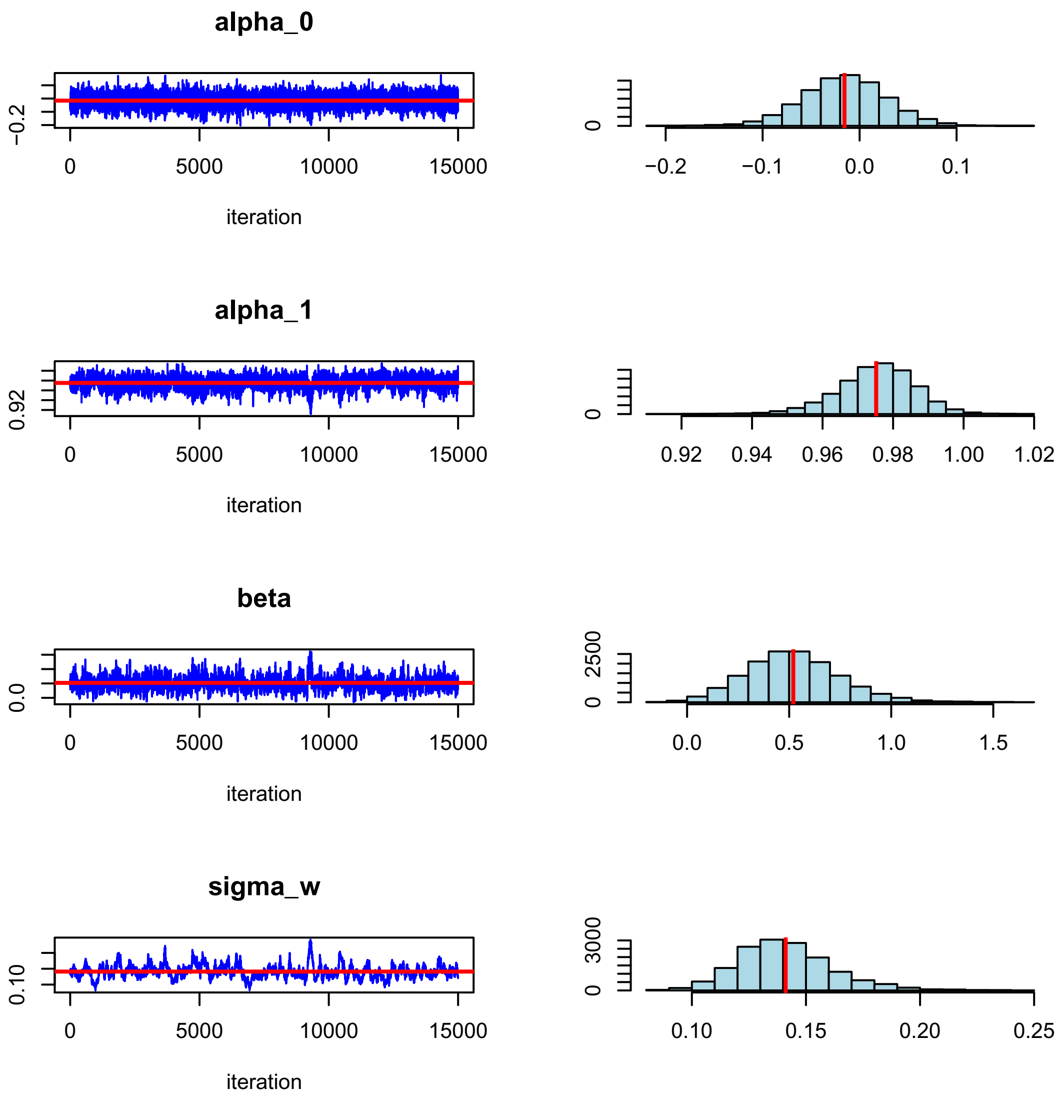

Figure 4: Results of Stochastic Volatility Model with Overnight Shocks

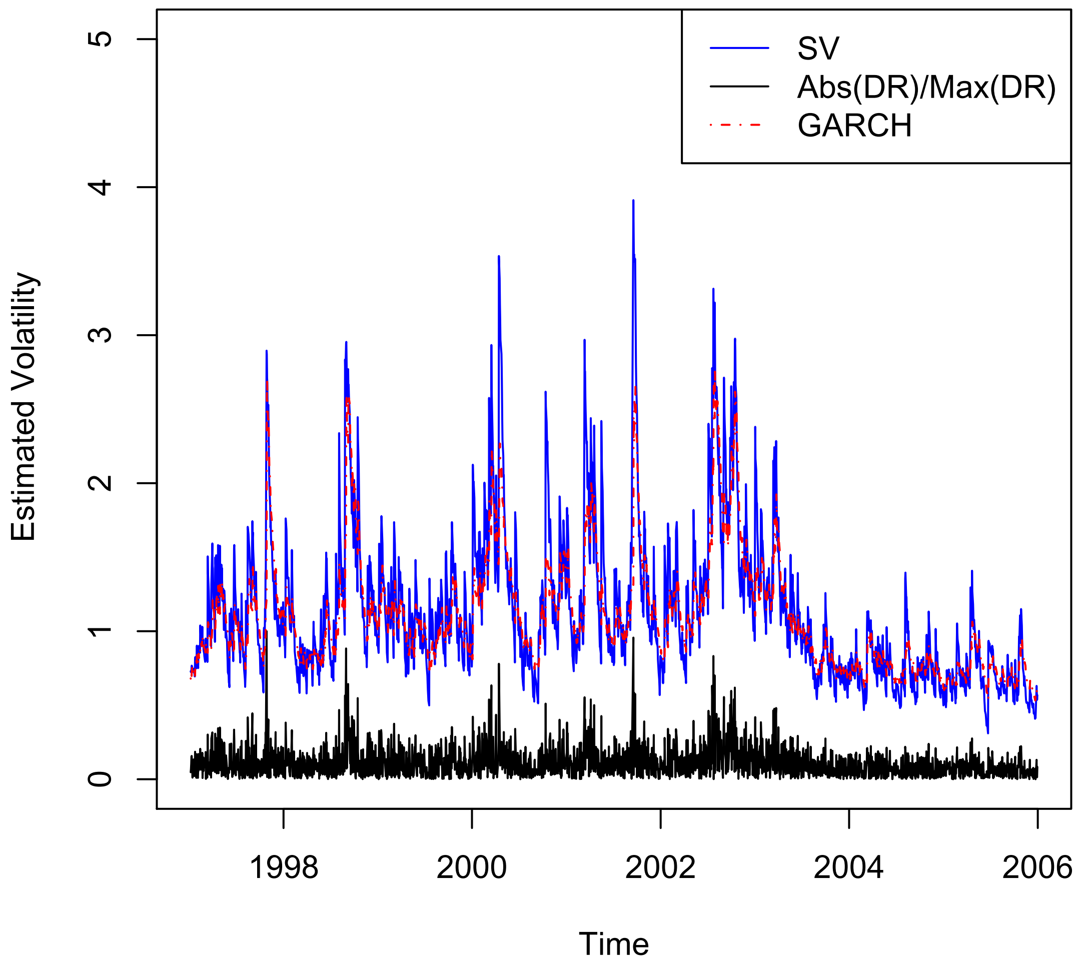

Figure 5: Volatility of SV and GARCH Model

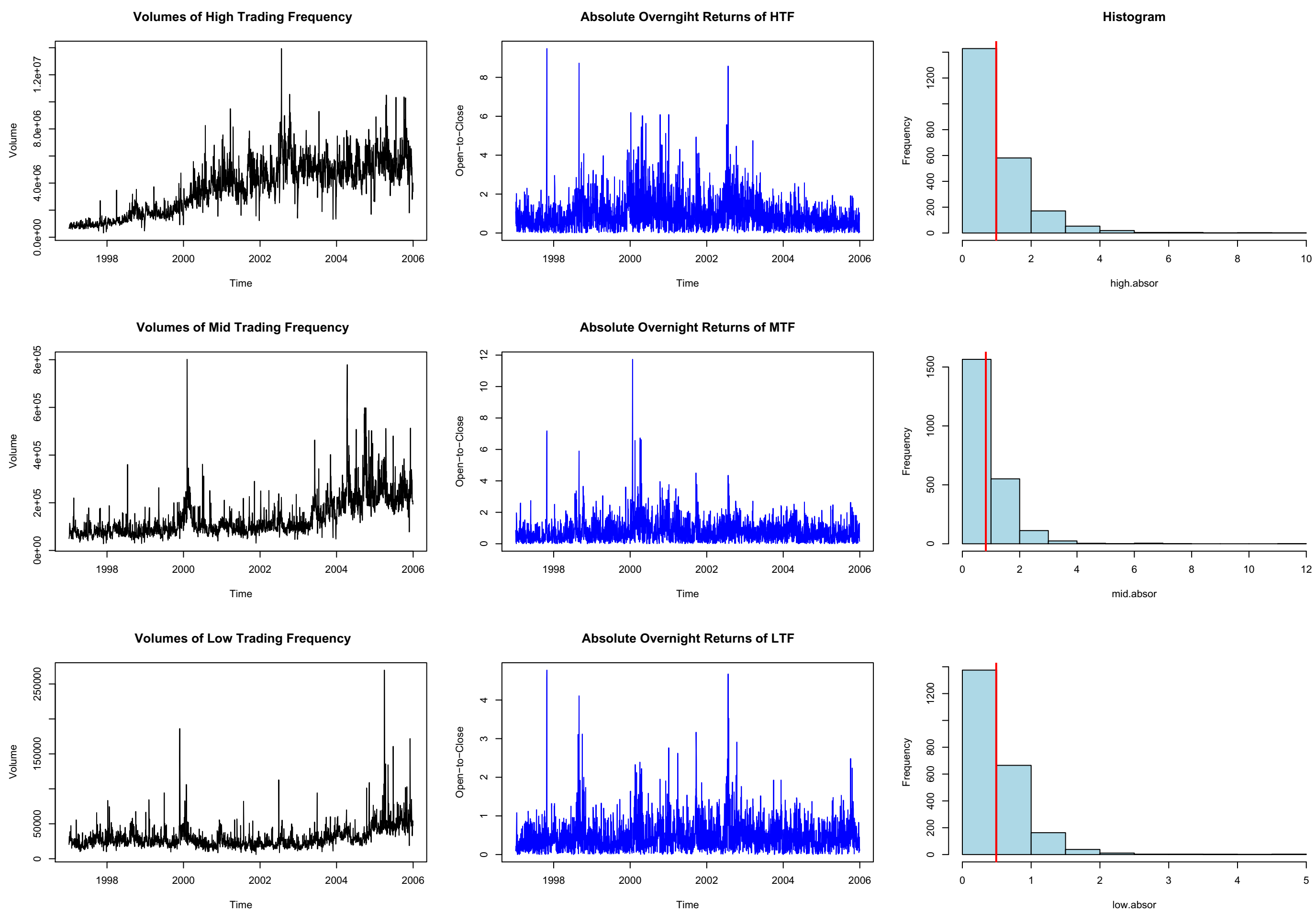

Figure 6: Summary for Different Trading Frequency Groups

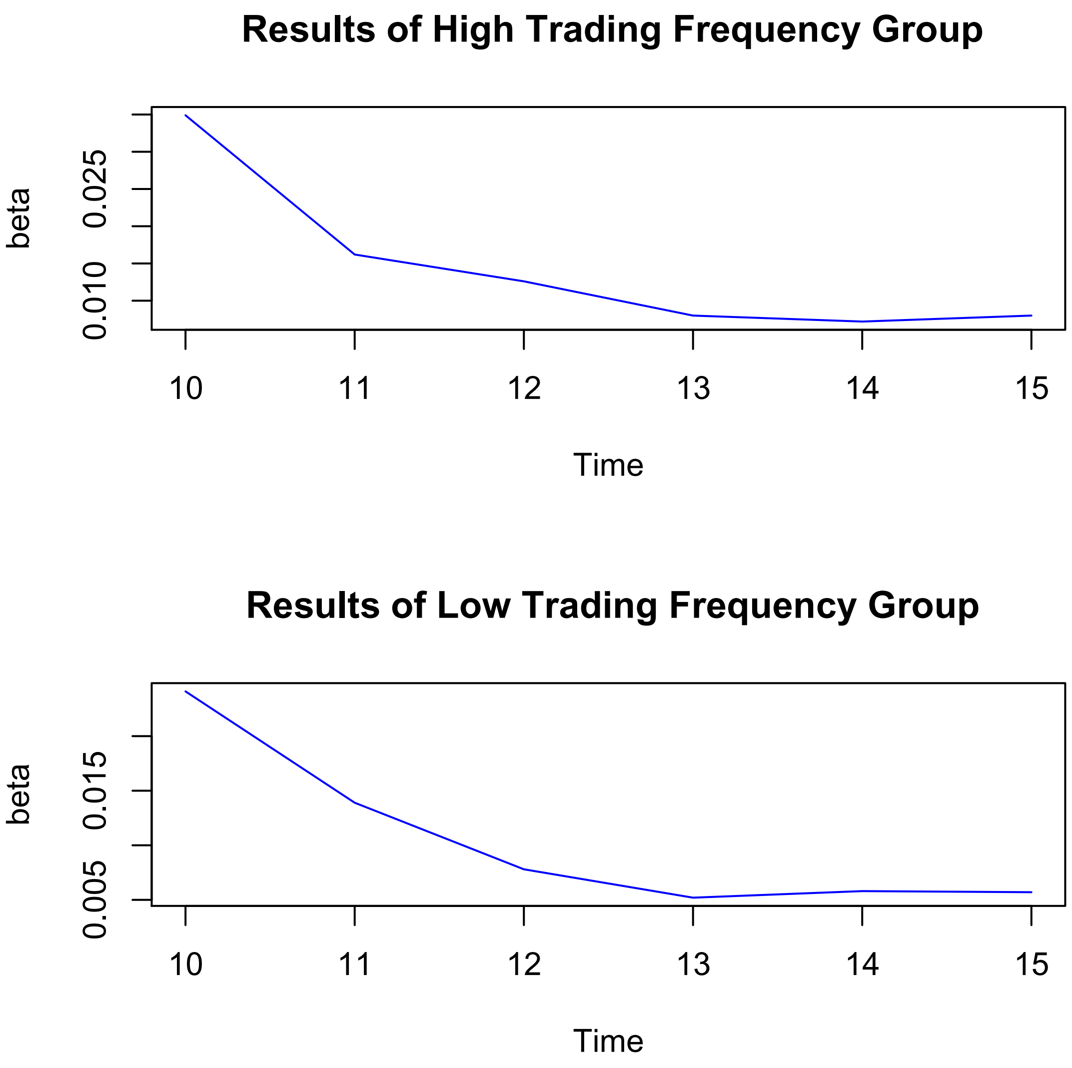

Figure 7: Intraday Estimates of β