

UBS Trading Insights: An Analysis of Daily and Intraday Trading Activities

ORIE 5640
Spring 2023
October 20, 2023

Yun Liu, yl2649
Nicolaj Hans Nielsen, nhn6

1 Introduction

On March 19th, UBS made an announcement regarding its plans to acquire Credit Suisse. This news was initiated by the Swiss Federal Department of Finance, FINMA, and the Swiss National Bank, making it a significant milestone not only for UBS and Credit Suisse but also for Switzerland and the global financial industry. In this project, we aim to analyze the trading activity of UBS and understand how this acquisition affects investor's trading strategies. Since such extreme events can have an immediate impact on the market, analyzing daily changes will not be sufficient to capture the full extent of the impact. Therefore, we will analyze both long-term and short-term impacts using time series analysis techniques and intraday trading data.

By combining these analyses, we will gain a comprehensive understanding of the overall dynamics of UBS' stock prices, especially after the acquisition and provide insights into the impact of significant financial events on trading activity in a leading investment bank.

2 Data and Methodology

2.1 Data Sources

To perform our data analysis, we require time series data at different frequencies for daily and intraday analysis. In both cases, we will model stock prices using log returns, which is a widely used approach in financial analysis and forecasting, but at different time scales: daily and every 5 minutes. Rather than modeling the stock prices directly, we focus on the percentage changes or returns of the stock prices over time as they have desirable statistical properties such as stationarity, which makes them easier to model and forecast [1].

2.1.1 Daily Data

We obtained the daily adjusted closing prices of UBS stocks from January 1st, 2015, to May 2nd, 2023, from Yahoo Finance. The data was imported using the quantmod package in R.

We then calculated the daily log-returns of the stock for our data analysis.

2.1.2 Intraday Data

For the intraday analysis, we acquired the 5-minute interval prices and volumes of UBS stocks from March 6th, 2023, at 9:30 AM to April 28th, 2023, at 3:55 PM. We calculated the log returns and used log volumes to analyze intraday trading activity.

2.2 Methodology

In our initial exploratory data analysis, we conducted KPSS and ADF tests on the time series of UBS 5-minute log returns and found that both tests indicated the time series was stationary. ACF and PACF plots (see Figure 33 in the appendix) and the Box-Pierce test showed no significant serial correlations in the 5-minute log returns. However, intraday data are usually too noisy [2], and additional data cleaning methods may be necessary. We also observed in Figure 5 that the intraday data was more volatile at the beginning of the day and gradually stabilized later on.

To capture these features, we decided to analyze the log returns of UBS stock at different time scales. As the intraday data did not exhibit significant correlation, we decomposed the intraday analysis into two independent directions: analyzing the 5-minute data cross days without considering the effect of different days, and analyzing the daily log returns. Our observation that the intraday data stabilized later in the day suggests that looking at the daily adjusted close price is a reasonable choice on a daily time scale since it is more stable and reflects the previous information of the day.

3 Data Analysis

3.1 Daily Trading Activity

This section presents the analysis results for daily log returns of UBS stock prices.

In our exploratory data analysis, we observed heavy-tailed distribution and volatility clustering behavior in the time series. To capture

the conditional heteroskedasticity and heavy tail behavior, we fitted various GARCH models. Apart from the standard GARCH model, we incorporated additional models to better capture the complexities of financial time series. Specifically, we fitted the GARCH-in-mean (GARCH-M) model, which adds a heteroskedasticity term into the mean equation to account for the dependence of a security's return on its volatility (risk) [3]. We also fitted the asymmetric power GARCH (apGARCH) model, which considers the phenomenon where large negative returns increase volatility more than positive returns of the same magnitude. To combine the risk premium and leveraging effect together, we also fitted apGARCH-M models. To better capture the heavy-tail and skewness behavior, we fitted normal, t , and skewed t -distribution for each model.

To compare the 12 fitted models, we examined various statistics and selected the model based on the Akaike Information Criteria (AIC). We also tested the model's forecasting performance using the method of cross-validation on a rolling basis, comparing the forecast of the 12 models with the test data of last 600 observations using Mean Square Error.

3.1.1 Exploratory Data Analysis

The time series exhibit the phenomena of volatility clustering and QQplots shows heavy tail behavior as suggested in Figure 1.

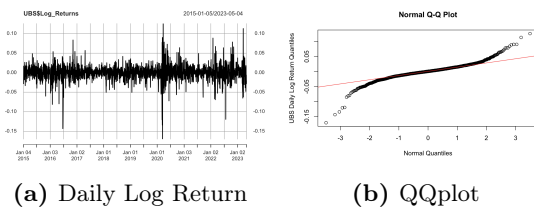


Figure 1 – Daily Log Returns of UBS

ACF plots in Figure 2 indicated no significant serial correlation in the log returns but showed correlation in the squared log returns, suggesting that GARCH models may be a good choice. Given the estimation difficulty and the success of GARCH(1,1) process for many financial time series [1], we utilize GARCH(1,1) process for all variations of GARCH process that we fitted.

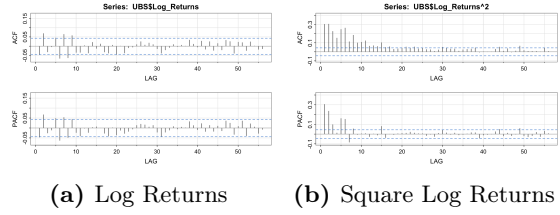


Figure 2 – ACF Plots of Daily Log Returns

3.1.2 GARCH Model Fitting Analysis

We began by fitting the standard GARCH model, which could not capture the heavy tail behavior according to the adjusted Pearson Goodness-of-Fit Test and QQ plots of the standardized residuals. The same issue occurred with the GARCH-M and apGARCH models, so we opted for fitting a heavy-tail distribution, ultimately choosing the t -distribution due to its widespread usage in financial time series.

After fitting the GARCH model with the t -distribution, we observed skewness in the QQ plots, prompting us to try fitting a skewed t -distribution. The QQ plots for the GARCH model with skewed t -distribution showed significant improvement compared to the previous models, with most of the statistical test results appearing reasonable, except for the Nyblom stability test for the parameter ω , which indicated a structural change within the time series. We proceeded to fit the GARCH-M and apGARCH models separately. For the GARCH-M model, the normal and t -distributions still exhibited the same issue as previously discussed, and while adding the skewed t -distribution made the p -value for the risk premium parameter more reasonable comparing to the other two distributions, the Nyblom stability test for the parameter ω did not improve.

For the apGARCH model, adding the leveraging effect improved the Nyblom stability test result for the parameter ω for all three distributions, although the normal distribution still failed to capture the heavy-tail behavior. The p -value of the parameter ω increased for both t and skewed t distributions so we cannot reject the hypothesis that $\omega = 0$, violating the GARCH model assumption. Therefore, adding only the leveraging effect was also not sufficient.

Motivated by the previous results, we decided to combine the models and consider the apGARCH-M model. Both the t and skewed t distribution models fit better than the previous models, and the skewness parameter is statistically significant. The AIC of apGARCH-M with the skewed t -distribution was smaller than the apGARCH-M with t -distribution, suggesting that the former was the better choice (in fact the AIC chose apGARCH-M with skewed t -distribution among all the models we fitted, see Table 2). Therefore, we chose the apGARCH-M with the skewed t -distribution as our final model to fit the daily log return of UBS. We plotted the empirical density, QQ plot, and ACF of standardized and squared standardized residuals in Figure 3 to demonstrate the fit of our model.

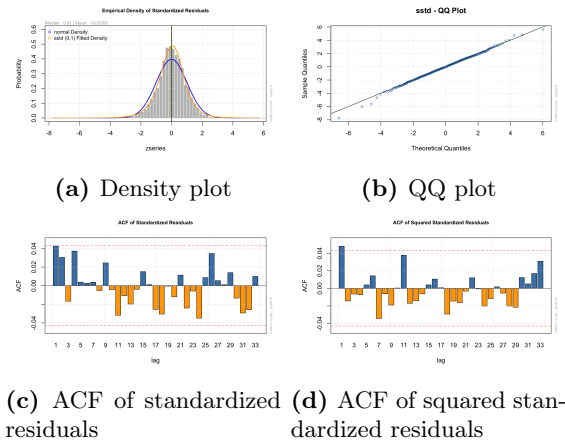


Figure 3 – Statistics for apGARCH-M model with skewed t -distribution

3.1.3 Model Validation

To test the efficiency of our model, we did a rolling window validation test for all 12 models on the the last 600 observations. Since updating the fitted model coefficients every time when we add one extra observation is very inefficient, instead we update the fitted parameters for every 50 new observations, and compare the forecasting results (see Figure 4) with the actual data and computed the mean square error (Table 3). It turns out that apGARCH-M with skewed t -distribution still performs the best among all the models we fitted.

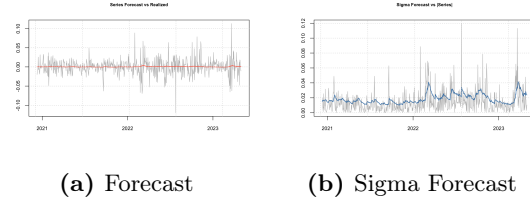


Figure 4 – Rolling window Validation for apGARCH-M with skewed t -distribution

3.2 Intraday Price Variation

This section aims to provide insight into when information is priced in throughout the trading day. Given the significance of the merger, our initial hypothesis was that unexpected information would cause significant movement in the stock price within the day. As shown in Figure 5, the log-returns for each 5-minute interval over the trading day reveal that the majority of the variation occurs within the first 5 minutes of the day.

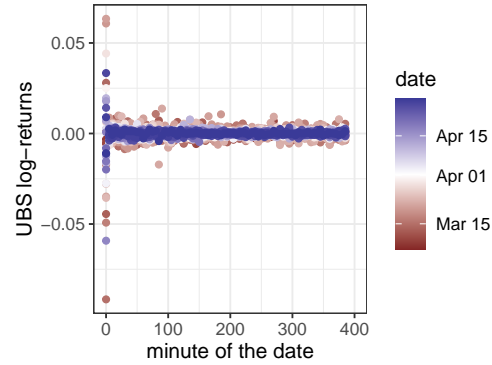


Figure 5 – Log-returns intraday over the period considered.

	PC1	PC2	PC3	PC4
Standard Deviation	0.031	0.006	0.006	0.005
Proportion of Var.	0.798	0.030	0.025	0.022
Cum. Proportion	0.798	0.828	0.852	0.874

Table 1 – Statistics for the first 4 components.

To gain a deeper understanding of the statistical variability in the data, we conducted a Principal Component Analysis (PCA) on the 5-minute data matrix. As displayed in Table 1, the first component accounts for nearly 80% of the total variation in the data, while the following components contribute only minor information in comparison. Upon examining Figure 6,

it is clear that the first eigenvector is heavily influenced by the first 5-minute interval.

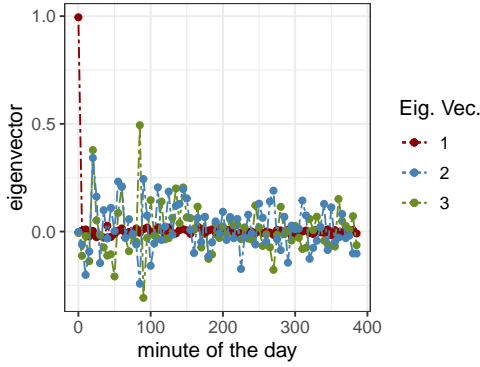


Figure 6 – The eigenvectors from the PCA.

Despite the significance of the merger, it appears that most of the relevant information was incorporated into the stock price within the first few minutes of the stock market’s opening. This finding is not uncommon, as other studies such as [4] have also reported similar patterns, which they attribute to the fact that most firm-specific information and earnings announcements occur outside of exchange opening hours. As a result, we have a large number of data points within a trading day that contribute little to the overall variation. This is one of the reasons why we employed daily log returns to analyze the long-term dynamics and period leading up to the merger. However, this finding raises the question of what traders are doing for the remainder of the day and which days around the merger experienced the most trading activity.

3.3 Intraday Trading Activity

This section aims to provide insight into the intraday distribution of trading volume during the period surrounding the announcement and aftermath of the merger, utilizing the realized 5-minute volumes. Our objective is to create a descriptive model of the volume during this period. Initially, we observed that a log-transformation was appropriate to handle the days with exceptionally high trading volume. The log-transformed data is depicted in Figure 7. The figure reveals that trading activity follows a U-shaped pattern, with higher activity at the beginning and end of the trading

day, a stylized characteristic of intraday trading activity [2].

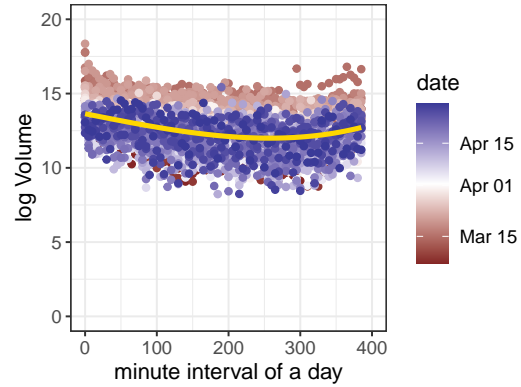


Figure 7 – Log-Volume intraday over the period considered. The line are fitted B-splines.

Let $Y_{d,m}$ denote the **log** volume on day d and intraday 5-minute m . We introduced a B-Spline Basis of polynomial splines to capture the U-shaped pattern of intraday trading volume. We accomplished this using **bs** in the package **spline** with **degree=3**. We then introduced a linear model,

$$Y_{d,m} = \alpha + \Psi_U(m) + \epsilon_{d,m} \quad (3.1)$$

where $\Psi_U(m) = \sum_{k=1}^3 a_k bs_k(m)$ is the sum of splines with fitted coefficient a_k , and $\epsilon_{d,m} \sim \mathcal{N}(0, \sigma^2)$. The found model coefficients can be seen in Appendix B.3.

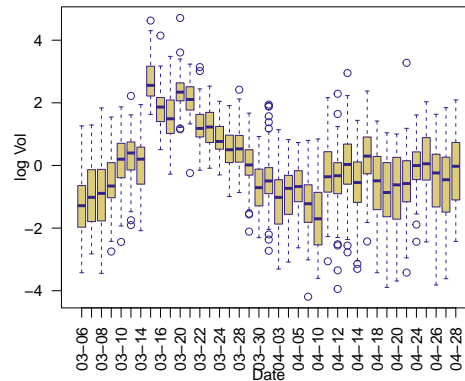


Figure 8 – Boxplot of the residuals for the model in Equation 3.1. All dates are in 2023.

The model appears to capture the primary intraday shape as seen in Figure 7. However, the residuals still show autocorrelation (see Figure Figure 34), and we observe that the mean and variance are different for each day in Figure Figure 8. Therefore, we need to account for this in

our model estimation. We will use the generalized least squares function, `gls`, from the `nlme` package. Inspired by the approach described in section 4.1.3 of [5], we will introduce a different variance parameter for each day. Our model is now given by:

$$Y_{d,m} = \alpha + \Psi_U(m) + \sum_{d=1}^N \beta_d \mathbb{1}_{\{d\}} + \epsilon_{d,m},$$

$$\epsilon_{d,m} \sim \mathcal{N}(0, \sigma^2 \Lambda_d) \quad \Lambda_d = \delta_d^2.$$

To ensure *identifiability*, we imposed the constraint $\delta_1 = 1$, making $\delta_2, \delta_3, \dots, \delta_N$ the ratios between the standard deviation of the first date and the subsequent dates, i.e. $\delta_j = \frac{\sigma_j^2}{\sigma_1^2}, j > 1$. To account for the intraday autocorrelation, we introduced an autocorrelation structure to our `gls` function, as suggested in examples in [6]. We found that an ARMA(2,1) correlation structure is adequate, and thus we modify our covariance matrix:

$$\Lambda_i = \delta_i^2 \begin{bmatrix} 1 & \rho(0) & \rho(2) & \cdots & \rho(n-1) \\ \rho(0) & 1 & \rho(0) & \cdots & \rho(n-2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho(n-1) & \rho(n-2) & \rho(n-3) & \cdots & 1 \end{bmatrix},$$

where $\rho(\cdot)$ is the autocorrelation function of an ARMA(2,1) process [7]. However, upon fitting the previous model, we observed clear indications of overparameterization in the variance structure. Therefore, we made modifications to the model such that we had the same ARMA structure within the day but shared variance parameters δ_i over the week. To examine the deficiencies of the model, we used normalized residuals, as defined in [6] section 5.3.4,

$$\mathbf{r}_i = \hat{\sigma}^{-1} \left(\hat{\Lambda}_i^{-1/2} \right)^T (\mathbf{y}_i - \hat{\mathbf{y}}_i).$$

We expect $\mathbf{r}_i \stackrel{iid}{\sim} \mathcal{N}(0, \mathbf{I})$, and we show this aligns with the graphs in Figure 35a and Figure 35b. What is important is that we have adequately parameterized the variance, as can be seen in Figure 9. The final model had more parameters, but the AIC and BIC were still much better. The initial model in Equation 3.1 had $AIC = 10584$ and $BIC = 10614$, while the final model reduced to $AIC = 7612$ and $BIC = 7930$.

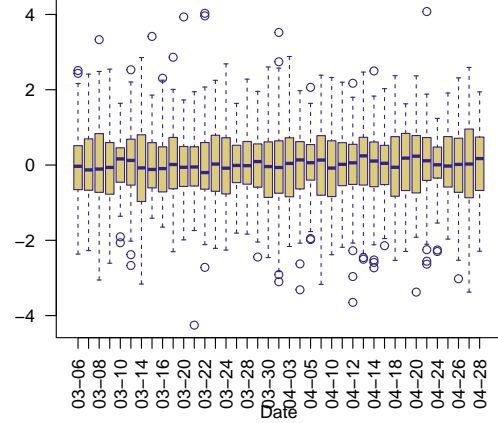


Figure 9 – Boxplot of the residuals for the final model. All dates are in 2023.

In the final model, we estimated the parameters and their 95% confidence intervals, as shown in Appendix B.5. Notably, not all days had an intercept significantly different from the common intercept α . However, all the days leading up to and following the 19th of March had intercepts significantly larger than the common intercept. Interestingly, the variance for the week following the merger showed much lower variation in the log volume. This suggests that activity was persistently high without many minutes of low activity during that period.

4 Conclusion

In conclusion, our analysis of the UBS stock data revealed interesting patterns in both the daily and intraday log returns.

For daily log returns, we found evidence of volatility clustering and heavy tail behavior and selected the apGARCH-M model with a skewed t -distribution to capture the risk premium and leveraging effect. This model performed well in both statistical tests and rolling window model validation.

For the intraday data, we found that most of the variation happens within the first 5 minute interval of the day. Regarding intraday volume, we observed a U-shaped pattern in the log volume throughout the day, with the highest activity at the start and end of the day. We also found that the days around the merger saw significantly larger volumes, with lower variation in the log volumes.

References

- [1] D. Ruppert and D. S. Matteson, *Statistics and data analysis for financial engineering*. Springer, 2011, vol. 13.
- [2] R. S. Tsay, *Analysis of financial time series*. John Wiley & Sons, 2005.
- [3] R. F. Engle, D. M. Lilien, and R. P. Robins, “Estimating time varying risk premia in the term structure: The arch-m model,” *Econometrica: journal of the Econometric Society*, pp. 391–407, 1987.
- [4] D. Lou, C. Polk, and S. Skouras, “A tug of war: Overnight versus intraday expected returns,” *Journal of Financial Economics*, vol. 134, no. 1, pp. 192–213, 2019.
- [5] A. F. Zuur, E. N. Ieno, N. J. Walker, A. A. Saveliev, G. M. Smith *et al.*, *Mixed effects models and extensions in ecology with R*. Springer, 2009, vol. 574.
- [6] J. Pinheiro and D. Bates, *Mixed-effects models in S and S-PLUS*. Springer science & business media, 2006.
- [7] R. H. Shumway, D. S. Stoffer, and D. S. Stoffer, *Time series analysis and its applications*. Springer, 2000, vol. 3.

A Daily

A.1 Model Fitting

The empirical density, QQ plot, and ACF of standardized and squared standardized residuals for the other 11 models are in Figure 10-20.

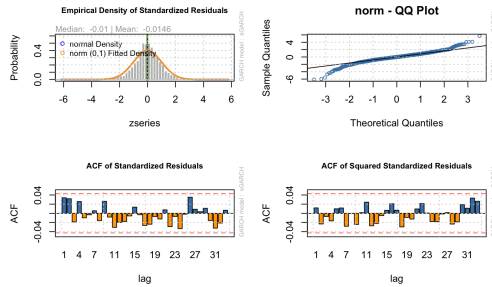


Figure 10 – standard GARCH model

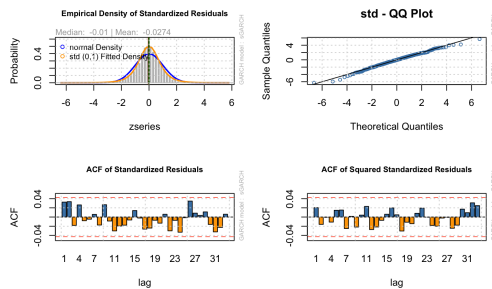


Figure 11 – GARCH model with t -distribution

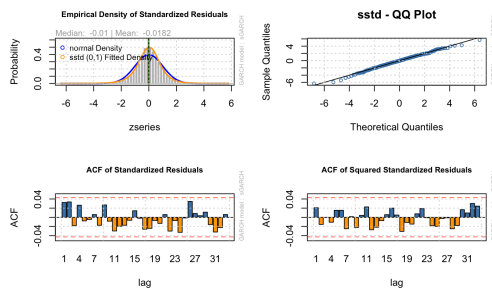


Figure 12 – GARCH model with skewed t -distribution

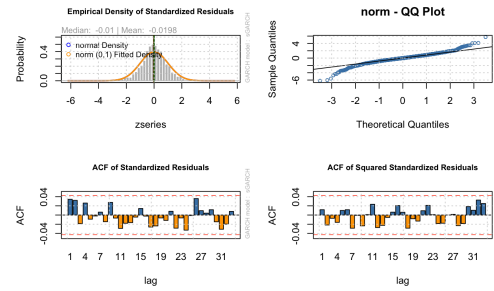


Figure 13 – GARCH-M model

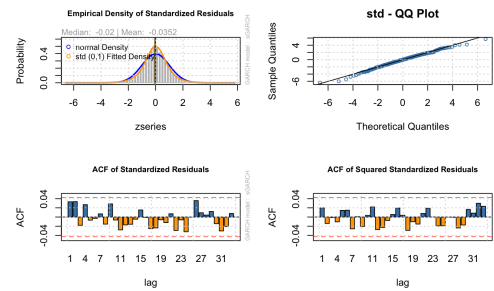


Figure 14 – GARCH-M model with t -distribution

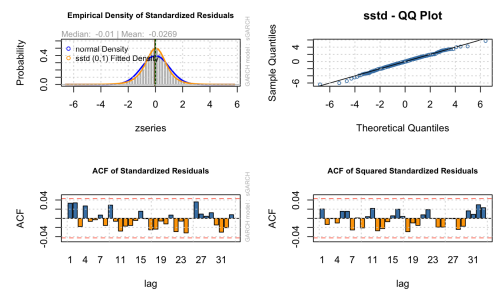


Figure 15 – GARCH-M model with skewed t -distribution

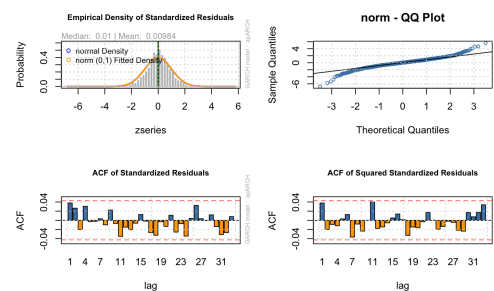


Figure 16 – apGARCH model

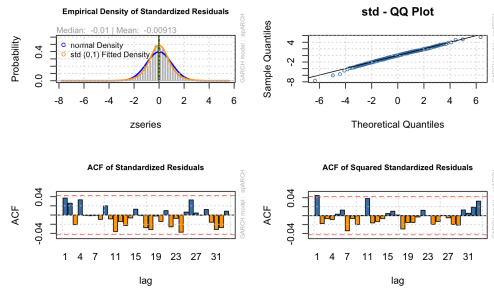


Figure 17 – apGARCH model with t -distribution

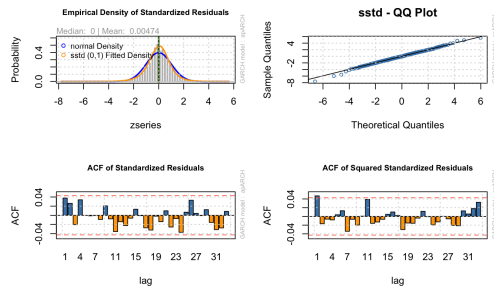


Figure 18 – apGARCH model with skewed t -distribution

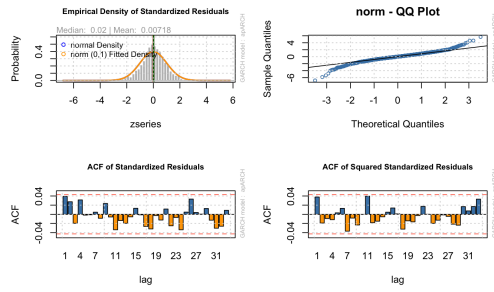


Figure 19 – apGARCH-M model

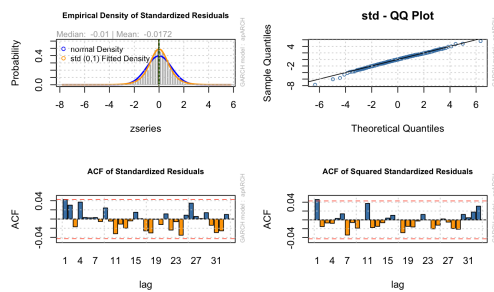


Figure 20 – apGARCH-M model with t -distribution

The AIC of the 12 fitted models are in Table 2.

Model	AIC
GARCH	-5.2942
GARCH-t	-5.3995
GARCH-st	-5.3991
GARCH-M	-5.2938
GARCH-M-t	-5.3999
GARCH-M-st	-5.3993
apGARCH	-5.3171
apGARCH-t	-5.4153
apGARCH-st	-5.4156
apGARCH-M	-5.3164
apGARHC-M-t	-5.4156
apGARCH-M-st	-5.4158

Table 2 – MSE of Model Validation

A.2 Rolling Window Model Validation

The forecast density, forecast sigma and forecast time series and Value at Risk (0.01) of the rolling window forecast of the 12 fitted model for the last 600 observations are in Figure 21-32.

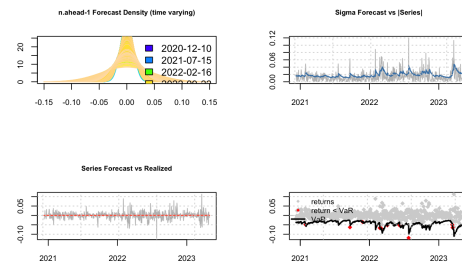


Figure 21 – Rolling window forecasting for standard GARCH model

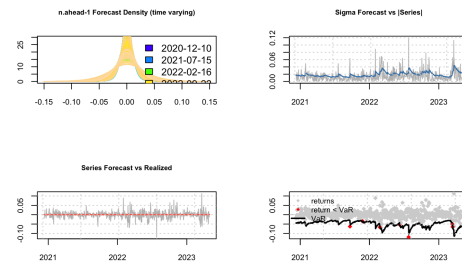


Figure 22 – Rolling window forecasting for GARCH model with t -distribution

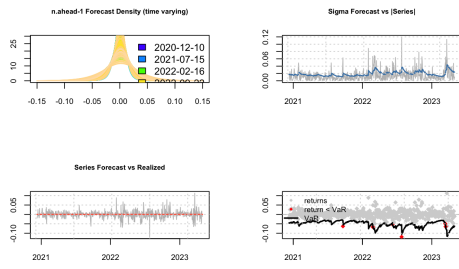


Figure 23 – Rolling window forecasting for GARCH model with skewed t -distribution

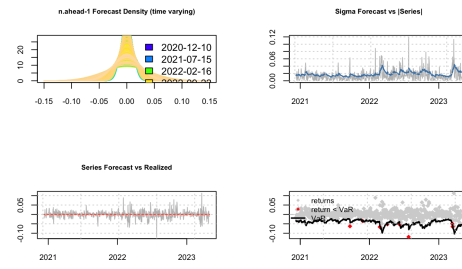


Figure 27 – Rolling window forecasting for apGARCH model

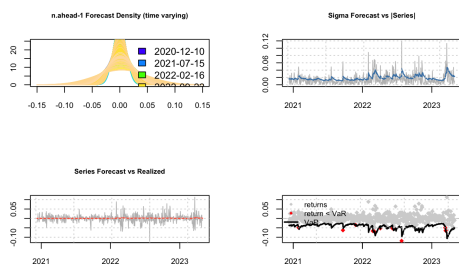


Figure 24 – Rolling window forecasting for GARCH-M model

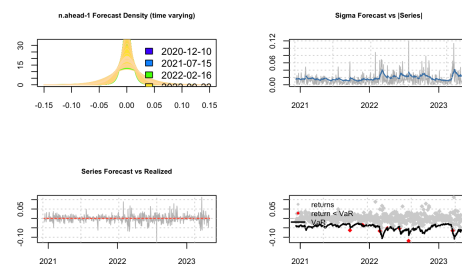


Figure 28 – Rolling window forecasting for apGARCH model with t -distribution

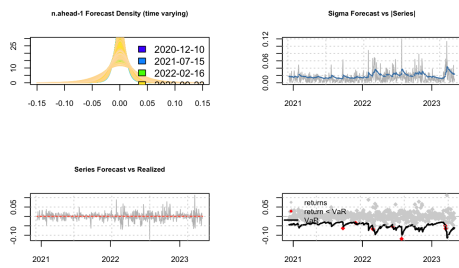


Figure 25 – Rolling window forecasting for GARCH-M model with t -distribution

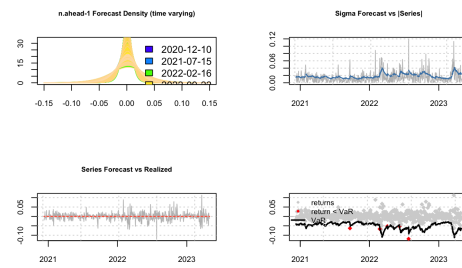


Figure 29 – Rolling window forecasting for apGARCH model with skewed t -distribution

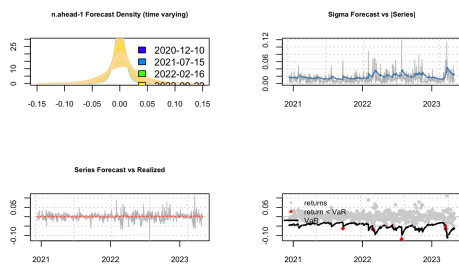


Figure 26 – Rolling window forecasting for GARCH-M model with skewed t -distribution

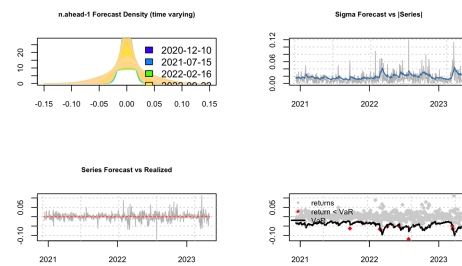


Figure 30 – Rolling window forecasting for apGARCH-M model

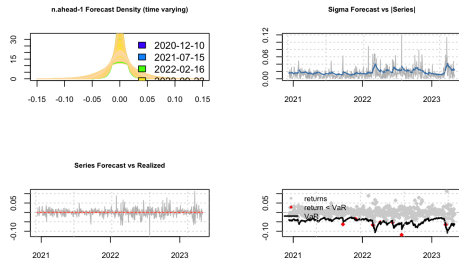


Figure 31 – Rolling window forecasting for apGARCH-M model with t -distribution

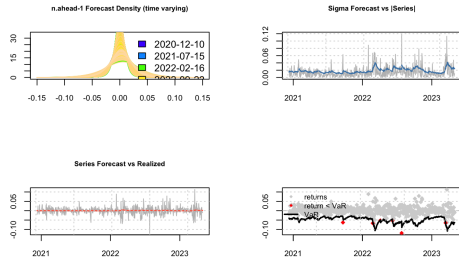


Figure 32 – Rolling window forecasting for apGARCH-M model with skewed t -distribution

The MSE, MAE and DAC of the rolling window forecast of 12 models are in Table 3.

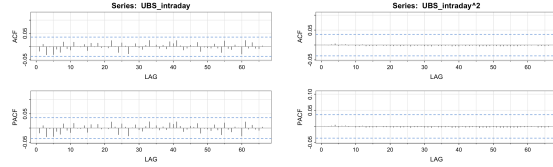
Model	MSE
GARCH	0.0004259993
GARCH-t	0.000425732
GARCH-st	0.0004258645
GARCH-M	0.0004251145
GARCH-M-t	0.0004248295
GARCH-M-st	0.0004248167
apGARCH	0.0004262429
apGARCH-t	0.0004258914
apGARCH-st	0.0004261361
apGARCH-M	0.0004251766
apGARHC-M-t	0.0004244092
apGARCH-M-st	0.0004243968

Table 3 – MSE of Model Validation

B Intraday

B.1 ACF Plots

The ACF plots of UBS intraday log return and squared intraday log return are in Figure 33.



(a) Log return

(b) Squared Log return

Figure 33 – ACF of UBS Intraday Log Return

B.2 Intraday Price Variation, PCA

	SD	Proportion of Var	Cumulative Proportion
PCA1	0.031	0.798	0.798
PCA2	0.006	0.030	0.828
PCA3	0.006	0.025	0.852
PCA4	0.005	0.022	0.874
PCA5	0.004	0.016	0.890
PCA6	0.004	0.015	0.905
PCA7	0.004	0.011	0.916
PCA8	0.004	0.011	0.927
PCA9	0.003	0.010	0.937
PCA10	0.003	0.007	0.944
PCA11	0.003	0.007	0.951
PCA12	0.003	0.005	0.956
PCA13	0.002	0.005	0.961
PCA14	0.002	0.004	0.965
PCA15	0.002	0.004	0.969
PCA16	0.002	0.003	0.972
PCA17	0.002	0.003	0.975
PCA18	0.002	0.003	0.978
PCA19	0.002	0.003	0.980
PCA20	0.002	0.002	0.983
PCA21	0.002	0.002	0.985
PCA22	0.002	0.002	0.987
PCA23	0.001	0.002	0.989
PCA24	0.001	0.002	0.990
PCA25	0.001	0.002	0.992
PCA26	0.001	0.001	0.993
PCA27	0.001	0.001	0.994
PCA28	0.001	0.001	0.995
PCA29	0.001	0.001	0.996
PCA30	0.001	0.001	0.997
PCA31	0.001	0.001	0.997
PCA32	0.001	0.001	0.998
PCA33	0.001	0.000	0.999
PCA34	0.001	0.000	0.999
PCA35	0.001	0.000	0.999
PCA36	0.001	0.000	1.000
PCA37	0.001	0.000	1.000
PCA38	0.000	0.000	1.000

Table 4 – All of the principal components and variance they explain.

B.3 Intraday Trading Activity, the Linear Model

	lower	est.	upper
(Intercept)	13.456	13.643	13.831
bs(min_of_day, df = 3)1	-1.768	-1.225	-0.681
bs(min_of_day, df = 3)2	-2.823	-2.471	-2.119
bs(min_of_day, df = 3)3	-1.221	-0.928	-0.635
σ	1.312	1.378	1.450

Table 5 – The parameters for the initial linear model with 95%-interval.

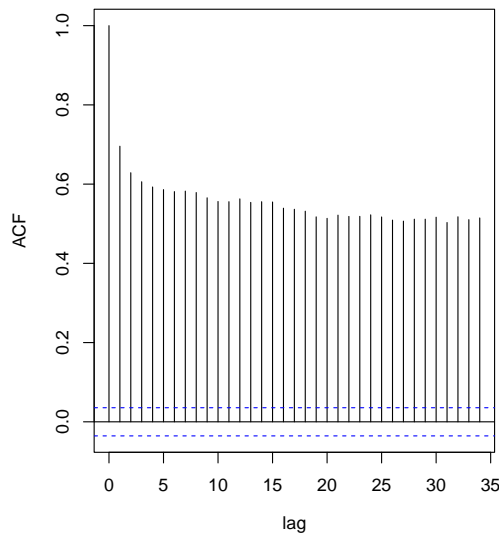
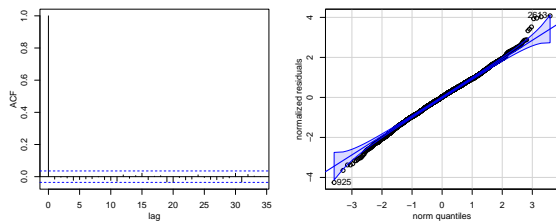


Figure 34 – The ACF of the linear model, clearly indicating autocorrelation.

B.4 Intraday Trading Activity, Residuals Diagnostics for the Final Model



(a) The ACF after fitting the final model for the log volume. (b) qqplot after fitting the final model for the log volume.

Figure 35 – Residual diagnostics for the final model.

B.5 Intraday Trading Activity, Parameters for the Final model

	lower	est.	upper
ϕ_1	1.068	1.109	1.088
ϕ_2	-0.236	-0.175	-0.113
θ_1	-0.892	-0.816	-0.695

Table 6 – Parameters for the ARMA process of the residuals.

week number	lower	est.	upper
11	0.614	0.671	0.733
12	0.475	0.519	0.568
13	0.667	0.729	0.796
14	0.769	0.847	0.932
15	0.900	0.975	1.057
16	0.928	1.012	1.103
17	0.926	1.020	1.123

Table 7 – δ_i variance parameters. Remember the first week has $\delta_i=1$ by construction.

	lower	est.	upper
(Intercept)	12.05	12.64	13.23
bs(min_of_day, df = 3)1	-2.04	-1.54	-1.04
bs(min_of_day, df = 3)2	-2.89	-2.49	-2.09
bs(min_of_day, df = 3)3	-1.32	-1.10	-0.89
2023-03-07	-0.58	0.16	0.89
2023-03-08	-0.52	0.26	1.04
2023-03-09	-0.29	0.49	1.28
2023-03-10	0.46	1.26	2.05
2023-03-13	0.69	1.37	2.04
2023-03-14	0.64	1.31	1.98
2023-03-15	3.15	3.83	4.50
2023-03-16	2.31	2.98	3.65
2023-03-17	2.03	2.71	3.38
2023-03-20	2.84	3.48	4.11
2023-03-21	2.63	3.26	3.89
2023-03-22	1.83	2.46	3.09
2023-03-23	1.70	2.34	2.97
2023-03-24	1.34	1.97	2.61
2023-03-27	0.94	1.63	2.33
2023-03-28	0.89	1.59	2.28
2023-03-29	0.42	1.11	1.80
2023-03-30	-0.20	0.49	1.18
2023-03-31	0.06	0.76	1.46
2023-04-03	-0.71	0.03	0.76
2023-04-04	-0.43	0.30	1.03
2023-04-05	-0.29	0.44	1.17
2023-04-06	-0.82	-0.08	0.65
2023-04-10	-1.25	-0.46	0.32
2023-04-11	-0.05	0.73	1.51
2023-04-12	-0.05	0.73	1.50
2023-04-13	0.15	0.92	1.70
2023-04-14	-0.06	0.73	1.51
2023-04-17	0.65	1.45	2.25
2023-04-18	-0.21	0.58	1.38
2023-04-19	-0.63	0.16	0.95
2023-04-20	-0.38	0.41	1.21
2023-04-21	-0.32	0.48	1.28
2023-04-24	0.39	1.20	2.00
2023-04-25	0.45	1.24	2.04
2023-04-26	0.00	0.80	1.59
2023-04-27	-0.40	0.40	1.19
2023-04-28	0.05	0.85	1.65

Table 8 – Intercepts with 95% confidence interval for the final model.