



Analyzing the Dynamic Relationship Between Intraday Trading Activity and Volatility Using High-Frequency Data

Emily Baker, Ryan Candy, and Isadora Coughlin

Advisors: Fangfang Wang and Jian Zou

March 29, 2021

A Major Qualifying Project report submitted to the
faculty of Worcester Polytechnic Institute in partial fulfillment
of the requirements for the degree of Bachelor of Science.

This report represents the work of WPI undergraduate students submitted to the faculty as evidence of completion of a degree requirement. WPI routinely publishes these reports on its website without editorial or peer review. For more information about the projects program at WPI, please see <http://www.wpi.edu/academics/ugradstudies/project-learning.html>.

Abstract

Measuring equity volatility is an important metric and understanding, describing, and predicting volatility using trading activity provides insight into navigating the stock market. Taking inspiration from existing research analyzing volatility in the stock market, we explore the dynamic relationship between trading volume, trading frequency, and volatility on an intraday basis across ten stocks in the consumer discretionary sector of the S&P 100 for the fourth quarter of 2013. Using three different volatility measures we implement variations of the heterogeneous autoregressive model and vector autoregressive model to investigate the lead-and-lag relationship between volatility and trading activity. Our quantitative analysis provides strong empirical evidence that current trading frequency and trading volume can be used to predict 30-minute measures of volatility and that the prior day rolling average and lagged trading measures are useful predictors in modeling the volatility measures.

Acknowledgements

We would like to thanks the following people who made this project possible. Without their help and assistance throughout this process, this project would not have been feasible.

First and foremost, thank you to our advisors, **Professors Fangfang Wang and Jian Zou** of the WPI Mathematical Sciences Department, for their guidance, advice, and feedback throughout the duration of our project.

We would also like to thank **James Kingsley**, Computational Scientist in the WPI Academic & Research Computing group at Worcester Polytechnic Institute, for providing consulting support that contributed to the results reported within this paper.

Additionally, we would like to thank **Professor Alexander Smith** of WPI's Social Science & Policy Studies Department for generously giving his time to answer our questions regarding the stock market, the New York Stock Exchange, and volatility in economics.

We want to thank **Yanzhao Wang**, a PhD student in WPI's Mathematical Sciences Department, for his advice and assistance in working with and cleaning the data from the Wharton Research Data Services database.

Finally, we want to thank **Leah Mitchell**, WPI Class of 2020, for her assistance in learning and writing in LaTeX.

Table of Contents

Table of Figures	v
Table of Tables	vi
1 Introduction	1
1.1 Literature Review	1
1.2 Research Questions	4
2 Data	5
2.1 Data Overview	5
2.1.1 Wharton Research Data Services	5
2.1.2 New York Stock Exchange and the S&P 100	5
2.2 Data Cleaning and Calculations of Variables of Interest	6
2.2.1 Realized Volatility	7
2.2.2 Range All and Five Minute Range	7
2.2.3 Summary Data	8
2.3 Exploratory Data Analysis	8
2.3.1 Trading Variables	8
2.3.2 Volatility Measures	10
2.3.3 Relationship between Trading and Volatility Variables	12
3 Heterogeneous Auto-Regressive Model	16
3.1 HAR Overview	16
3.2 Implementation	17
3.3 Models	17
3.4 Results and Analysis	18
4 Vector Auto-Regressive Model	21
4.1 VAR Overview	21
4.2 Implementation	21
4.3 Models	22
4.4 Results and Analysis	22
4.4.1 Predicting Volatility Measures	23
4.4.2 Predicting Trading Measures	24
5 Conclusions	25

6	Next Steps	27
6.1	Alternative Data Cleaning Methods	27
6.2	Expanding the Time Frame	27
6.3	Dividing the Model	28
6.4	Examining a Change in TF/TV Covariate	28
6.5	Forecasting	28
	Bibliography	29
A	S&P 100 Stocks	30
A.1	Full List of Stocks in the S&P 100	30
A.2	Consumer Discretionary Stocks	32
B	Full HAR Results	33
C	Full VAR Results	37

Table of Figures

2.1	Time Series of Trading Frequency	8
2.2	Day by Day Time Series for Trading Frequency	9
2.3	Autocorrelation Plot for Target's Trading Volume	9
2.4	Numerical Summaries of Five Minute Range Data for All Stocks	10
2.5	Volatility Time Series for Target	11
2.6	Day by Day Time Series for Target	12
2.7	Autocorrelation Plot for Target's Five Minute Range Volatility	12
2.8	Summary Data Correlation Matrix	13
2.9	Trading Frequency vs. Realized Volatility Day by Day Time Series for Target	14
2.10	Ford CCF Plot Between Trading Volume and Realized Volatility	14
2.11	Trading Frequency vs. Realized Volatility Day by Day Time Series for Ford	14
2.12	CCF plot for Target between trading volume and realized volatility	15
3.1	HAR Covariate Significance and Model Fit for Target's Range All Volatility	18
4.1	VAR Covariate Significance and Model Fit for Target's Range All Volatility	23

Table of Tables

2.1	Numerical Summaries of Five Minute Range Data for All Stocks	10
3.1	HAR Covariate Coefficients for Target's Range All Volatility	19
4.1	Matrix A for Model VAR-TF	23
4.2	Matrix A for Model VAR-TV	23
6.1	Morning and Afternoon Split Model Results	28

Chapter 1

Introduction

The relationship between trading frequency, trading volume, and volatility is very complex. This chapter introduces the concepts and reviews the literature on how various researchers have explored the intricacies of this dynamic relationship. We sat down with Associate Professor of Economics at Worcester Polytechnic Institute, Alexander Smith, to further understand exactly what volatility in the stock market is before we began our own exploration of the relationship.

Professor Smith explains that volatility is “all about the ups and downs,” and more formally, “a measure of how much movement you have around a trend.” Volatility is a good indicator of how confident people are in a particular stock’s future prospects. For example, if a stock’s price is constantly fluctuating with no comprehensive pattern, then the price is very volatile. This would, as Smith explains, “make people nervous” about purchasing that particular stock. Conversely, if a stock is more consistent in its price, it is less volatile and people will be more likely to trade. This affects the trading volume of a stock, or the number of shares traded in a given period of time, as well as the trading frequency of the stock, or the number of transactions in a given period of time. As such, the relationship between trading volume, trading frequency, and volatility is complex and important to observe in order to understand how volatility is affected by trading measures.

The following section explains the various approaches researchers take to quantify the relationships between trading frequency, trading volume, and volatility in various markets. It is important to note that volatility is latent and there is not one accurate way to measure volatility and thus any conclusions are specific to the volatility measure used. Then, we describe the research questions for the project based on the literature we reviewed.

1.1 Literature Review

In their work, Cheriyan and Lazar (2019) examine a set of market stocks from India. They analyze volatility based on the stock’s liquidity and trading activity (volume and frequency). According to the paper, liquidity “indicates the occurrence of larger volumes of trade within the shortest possible time and with the least possible cost of transaction.” Based on their prior research of the market microstructure theory, they expect that a more liquid market will have a higher volume of trades and therefore will be less volatile. Their hypothesis predicts an inverse relationship between liquidity and volatility, a positive relationship between trading volume and volatility, and an inverse relationship between trading volume and liquidity.

Cheriyan and Lazar (2019) use intraday data in the form of one minute trade and quote returns from the 50 stocks in the NIFTY 50 Index. The data was collected from January 1st, 2016 to

December 31st, 2016, resulting in 246 trading days. They looked at a time span between 9:16AM and 3:29PM Indian Standard Time, creating 374 one minute intervals resulting in a total of 92,004 observations per stock. The study created nine different intraday, minute-by-minute liquidity measures. They calculated their volatility approximation by averaging the one minute returns (to create an average one minute return) across the 50 stocks to create the intraday market return. This information was brought to a generalized autoregressive conditional heteroskedasticity (GARCH) model with parameters (1,1) and the residuals were defined as the market volatility.

In order to analyze the relationship between liquidity, trading volume, and volatility, the researchers utilized ordinary least squares models. Their hypothesis of the inverse relationship between trading volume and liquidity was supported by their research. Their hypothesis of the positive relationship between trading volume and volatility was also supported with the coefficients expressing a positive, statistically significant relationship. The final component of their hypothesis, the inverse relationship between volatility and liquidity was not formally supported, but their research found a positive correlation between volatility and illiquidity, which implies that there is an inverse relationship between volatility and liquidity (Cheriyian and Lazar, 2019).

Fulvio Corsi (2009) examines the importance of choosing the correct model type in his paper *A Simple Approximate Long-Memory Model of Realized Volatility*. Since financial data can be sporadic, with large changes in price over short periods of time, models like standard GARCH and stochastic models of volatility struggle to reproduce these patterns as anything other than white noise. As such, it is important to consider a long-memory process. A method previously used to achieve this is fractional integration, which although it is a straightforward way to obtain long-memory volatility, there are many limitations on the calculations, especially mixing short term and long term characteristics. To try to counteract this, Corsi (2009) proposes the Heterogeneous Autoregressive model for Realized Volatility (HAR-RV). This model allows the user to take many aspects of financial modeling into account, such as different types of traders, and then observe the inter-relationships of volatility and draw conclusions.

To utilize these models, Corsi (2009) creates three time periods over which to observe the data: one day, one week, and one month. The model uses a cascade approach and assumes that each new model is a function of past realized volatility experienced at the same time scale. This is a similar philosophy to the economic theory that “each volatility component in the cascade corresponds to a market component that forms expectations for the next period’s volatility based on the observation of the current realized volatility and on the expectation for the longer horizon volatility” (Corsi, 2009).

Although this model is simple, it is effective at modeling realized volatility through previous price data over extended periods of time. To analyze the efficacy of the simulated HAR-RV model the results were compared to a long tick-by-tick USD/CHF series over a 14 year time period from December 1989 to December 2003. While the simulated data does have some smoothing effect, it follows similar patterns to the actual data, therefore showing that the HAR-RV model is an accurate simulation of financial data over a long period of time. The models were also used for one day ahead forecasting, where they performed accurately (Corsi, 2009).

Lu and Lin (2010) examine the relationship between intraday volatility and trading volume using a bivariate vector autoregressive model. Their paper explored four existing hypotheses around the volume-volatility relationship: the mixture of distributions hypothesis (MDH), the sequential arrival of information hypothesis (SAIH), the dispersion of beliefs hypothesis, and the noise trader hypothesis.

They examined data from the Taiwan stock exchange. Since volatility cannot be exactly measured, they approximated volatility using a calculated sum of intraday one-minute returns, or realized volatility. In their methodology, Lu and Lin (2010) use the vector autoregressive (VAR) model to

examine the more specific bidirectionality of the relationship. They determined the optimal order of the VAR model to be two. After running the VAR(2) model, they tested for the direction of Granger causality using an F-test.

Their research concludes that there is a very strong relationship between past trading volume and volatility and the current trading volume or volatility. More specifically, they explain that an increase of one standard deviation in trading volume is eventually followed by an increase in volatility, which peaks after 17 days. They find that trading volume reacts more “sluggishly,” as an increase of one standard deviation in volatility is followed by an increase in trading volume that peaks after three days. Their conclusions support the sequential arrival of information hypothesis (SAIH), but they also note that there was sample size limitations, and as such more information is required to extrapolate this information outside of their research (Lu and Lin, 2010).

On the volatility–volume relationship in energy futures markets using intraday data by Julien Chevallier and Benoît Sevi (2012) examines the relationship between trading volume and price volatility in the crude oil and natural gas futures markets when using high-frequency intraday data. This mirrors one of the central relationships examined in this project of regressing trading volume and frequency with realized volatility. Chevallier and Sevi (2012) focus on the energy futures market, and specifically on crude oil and natural gas, which are the two most liquid energy markets in the world. This provided the inspiration for this project to focus on one sector of stocks, instead of a broader population such as the S&P 100.

Chevallier and Sevi (2012) note that microstructure noise can decrease the ability of the predictors to actually capture and convey the trends they represent. Thus, predictors were sampled using five minute intervals. Similarly, realized volatility was calculated using the difference of five minute returns and finding the summation of their squares. They comment that it is “useful to use lagged values of trading volume to predict return volatility” but go on to say that overnight returns have no additional explanatory power on realized volatility when added to the predictive power of trading data from the current day.

This paper uses heteroskedasticity-robust standard errors (using Newey-West correction) when analyzing regressions to avoid the effects of collinearity in the results among the predictors (Chevallier and Sévi, 2012). Chevallier and Sevi (2012) found that there was a significant relationship between both trading volume and trading frequency with realized volatility, along with a relationship between all three in the same model. Although trading volume was found to be even more closely related, both predictors were significant at the $\alpha = .01$ level. These relationships were strongly positive and both trading volume and frequency shared approximately the same information content about the volatility measure. The authors also found a significant predictor that acted as a dummy variable for the day of the week that press releases related to the oil and gas industries were released. Not only did this variable improve the regression models, but it also shows that the focus on one industry allowed real-world events to be more carefully considered.

1.2 Research Questions

After reviewing literature related to the relationship between trading volume and frequency and volatility measures, we refine the scope of the project and will base our analysis on the following research questions:

1. **What is the dynamic relationship between trading volume, trading frequency, and volatility at the intraday level?**
2. **What models can effectively summarize and explain this relationship?**

As we will outline in Chapter 2, we originally considered intraday trading data from the stocks in the S&P 100 for the year of 2013. Ultimately, we analyzed the consumer discretionary stocks from the S&P 100 to focus on one sector of stocks in the fourth quarter of 2013, resulting in 62 trading days. We then outline two sets of models used to explain the trading measure - volatility relationship, one using heterogeneous autoregressive (HAR) models and the other using vector autoregressive (VAR) models in Chapters 3 and 4, respectively. Finally, we draw important conclusions and answer the above research questions in Chapter 5 and discuss potential future work, such as forecasting this relationship, in Chapter 6.

Chapter 2

Data

In this chapter, we will describe where the data set for this project was obtained, how the data was cleaned, and how we calculated the variables of interest. We will also give an overview of our exploratory data analysis, including an examination of trading volume, trading frequency, and volatility measures.

2.1 Data Overview

2.1.1 Wharton Research Data Services

We obtained our data through the Wharton Research Data Services (WRDS) database hosted by the Wharton School of Business at the University of Pennsylvania. This database contains data on a wide range of topics mostly related to the stock market. The section of the database we utilized for our project was the TAQ, or Trades and Quotes, section, and specifically, the millisecond consolidated trades. There are data available for the years 2009 through 2013. Users can select the time frame of data they are looking for, which for us was October 1, 2013 through December 31, 2013 and 9:30 am to 4:00 pm. When creating our queries we selected all 14 possible variables which are: date of trade, time of trade, security symbol root, security symbol suffix, exchange that issued the trade, trade sale condition, volume of trade, price of trade, trade stop stock indicator, trade correction indicator, trade sequence number, source of trade, trade reporting facility, and trade ID. When selecting our stocks of interest we started with all stocks on the S&P 100 in the year 2013. This resulted in a total of 101 stocks, since one company has two classes of stocks, which are listed in full in [Appendix A](#).

2.1.2 New York Stock Exchange and the S&P 100

The S&P 100 is a stock market index “designed to measure the performance of large-cap companies in the United States” ([S&P Dow Jones Indices, 2021](#)). Every stock in the S&P 100 is traded on the New York Stock Exchange (NYSE), which is “the largest securities exchange in the world” and “provides a platform for buying and selling over nine million corporate stocks and securities a day” ([Corporate Finance Institute, 2021](#)). Since we were looking at the S&P 100, a subset of the S&P 500, we were able to only analyze stocks traded on the NYSE. Although we did originally download data for all 101 stocks, as discussed above, we narrowed down our focus to the 10 stocks in the consumer discretionary sector, which is made up of “industries that tend to be the most sensitive

to economic cycles” (Fidelity, 2021). The final stocks we analyzed were Amazon (AMZN), Disney (DIS), Ford (F), Home Depot (HD), Lowe’s (LOW), McDonald’s (MCD), Nike (NKE), Starbucks (SBUX), Target (TGT), and Time Warner (TWX). During the data cleaning process we created our variables of interest, including realized volatility, five minute range, range all, trading frequency, and trading volume. Realized volatility, five minute range, and range all are all functions of a stock’s price, which we will talk about in more detail below.

2.2 Data Cleaning and Calculations of Variables of Interest

The data from WRDS originally contained both a lot of microstructure noise and additional information that was not relevant to our project. Subsequently, we created a cleaning procedure that isolated the data that is desired for our models. We first filtered the data using the following steps from *Realised Kernels in Practice: Trades and Quotes* (Barndorff-Nielsen et al., 2009):

1. Delete entries with a timestamp outside of the 9:30 am to 4 pm window when the NYSE is open
2. Delete entries with a bid, ask, or transaction price equal to zero
3. Delete entries not originating from the NYSE
4. Delete entries with corrected trades (where the correction indicator, *CORR*, was not equal to zero)
5. Delete entries with abnormal sale condition (where the sale condition, *COND*, is not “E” or “F”, representing an Automatic Execution or an Intermarket Sweep Order)
6. If multiple transactions have the same time stamp, replace each price in the set with the median price of the set

As noted in Chapter 1.1, there are significant amounts of microstructure noise that can adversely affect the analysis of the data. Thus, we computed prices over five minute intervals, where we selected the price from the timestamp at or immediately prior to each five minute mark throughout the trading day, which is commonly known as the previous tick method. For the 9:30 am time mark at the beginning of the trading day, we selected the price from the first trade of the day.

After cleaning the data, we then proceed to calculate the trading volume and frequency within each 30-minute interval of the trading day. This means there were 13 intervals per day, where the first 30 minute interval is [9:30 am to 10:00 am], the second is (10:00 am to 10:30 am], and so on until the final interval from (3:30 pm to 4:00 pm]. Trading volume was calculated by finding the summation of the number of shares traded in the given interval. Trading frequency was calculated by counting the number of transactions that occurred in the given interval.

Then, we calculated three volatility measures for each stock, realized volatility (RV), range all (RA), and five minute range (R5), which are defined in Sections 2.2.1 and 2.2.2 in this chapter. Once these measures were calculated, we were able to export a 13 row matrix for each trading day where the following variables were included:

- Month (integer representing the month)
- Date (integer representing the day number of the month)
- Weekday (string, M/T/W/R/F, representing the day of the week)

- Index (integer, representing the 30 minute interval the data corresponds to)
- Day Number (integer, unique identifier for each trading day in our time period of interest)
- Realized Volatility (real number, representing a volatility measure)
- Range All (real number, representing a volatility measure)
- Five Minute Range (real number, representing a volatility measure)
- Trading Volume (real number, representing a trading measure)
- Trading Frequency (real number, representing a trading measure)

2.2.1 Realized Volatility

First, to eliminate the impact of microstructure noise, the five min log returns of the prices were calculated using the following equation, where r is the five minute log return and $P(t)$ is the equity price at time t . The log return at time t , $r(t)$, is calculated as

$$r(t) = \log(P(t)) - \log(P(t-1)).$$

Then, to provide a set of measures for use in analysis and modeling for each trading day, realized volatility was calculated for each of the thirteen 30 minute intervals throughout the day (same as used for trading volume and frequency). Realized volatility was calculated by finding the summation of the squares of the five minute log returns within the interval (I), as seen below.

$$RV = \sum_{t \in I} r(t)^2.$$

2.2.2 Range All and Five Minute Range

We utilized two calculations for range to illustrate the trade off between the increased microstructure noise in the range all data with the reduced information conveyed by the five minute range data. It is interesting to note the comparison between these distributions in Chapter 2.3, where both measures follow similar distributions.

The full high-low range was calculated by finding the logarithmic difference between the maximum and minimum prices in each 30 minute interval. In other words, no five minute prices were used, instead all of the cleaned tick data was evaluated to find the minimum and maximum for each interval. The equation is noted below, where RA is the range all measure, and P is the price.

$$RA = \log(P_{max}) - \log(P_{min}).$$

The five minute high-low range was calculated by finding the logarithmic difference between the maximum and minimum prices among the five minute price data in each 30 minute interval. The equation is noted below, where $R5$ represents the five minute range measure, and P is the price. The important distinction is the maximum and minimum are evaluated from the five minute price data, not all of the cleaned tick data in the given interval, which distinguishes this measure from the range all measure.

$$R5 = \log(P_{max}) - \log(P_{min}).$$

2.2.3 Summary Data

In addition to calculating volatility and trading measures for individual stocks, we also calculated these measures for our summary data. To do this we took the median at each time stamp across all stocks for the calculated volatility measures and current trading frequency and volume. We are then able to run our models on data that encompasses the full picture. We chose to do this in order to have a set of data that adequately summarized information across all stocks regardless of inconsistencies and to aid us in seeing overall trends.

2.3 Exploratory Data Analysis

2.3.1 Trading Variables

Trading volume and frequency have extreme upper outliers across all stocks, illustrating the range of actions possible in stock market trading, which causes all 10 stocks to have a skewed right distribution. However, Ford's maximum level of trading volume was over 7 times greater than the second largest maximum, and their maximum level of trading frequency was almost double the second largest maximum. Similarly, Ford has the largest mean and median for trading volume. However, Disney has the largest mean and median for trading frequency. In both trading variables, Starbucks and Amazon have the smallest means and medians.

When looking at the time series plots for trading volume and frequency we can observe a fairly consistent pattern between stocks. While some stocks experienced more volatility than others, each stock shows a single spike in trading volume relative to the rest of the stock's data. Although each spike occurs at a different point in time, most likely due to individual large events in the company's history, they almost all have exactly one. Of all of the stocks, Ford experienced the largest spike, with nearly 5000 trades in a 30 minute interval, while all others had a largest spike between 1200 and 2500 trades. This specific spike can potentially be attributed to the company's announcement at an investor meeting that they would move forward with an aggressive plan to increase product launches the following year that would drive down profits ([Isidore, 2013](#)). We can see the difference between a typical time series plot as seen in Figure 2.1, represented by Target, and the Ford time series plot.

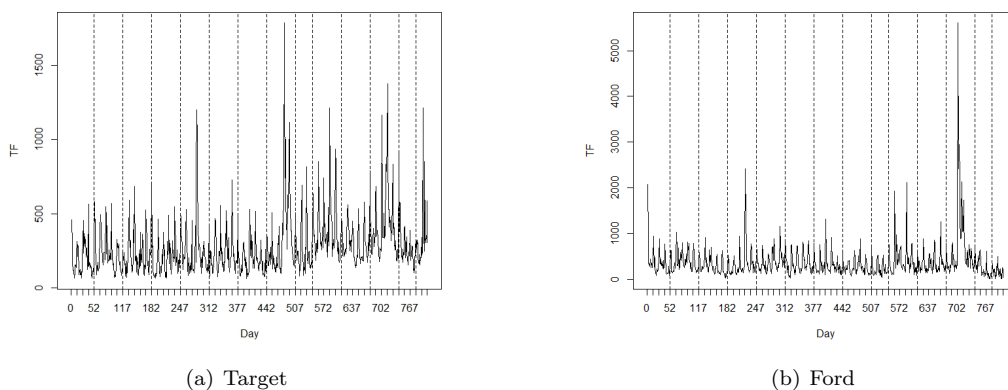


Figure 2.1: Time Series of Trading Frequency

The daily time series were calculated using the normalized average value for the trading frequency or volume during each 30 minute increment for a given weekday in order to observe the diurnal patterns in the data. As explained in the data cleaning process, we split up the 9:30 am 4:00 pm workday into thirteen 30 minute increments. According to [Mitchell \(2020\)](#) it was common practice for the lunch break to occur between 11:45 am and 1:30 pm, which is denoted by the vertical dotted lines in Figure 2.2. Both trading frequency and volume followed a similar daily pattern, with the patterns reaching a maximum at either the beginning of the workday and end of the workday, with a minimum around the end of lunchtime. It was also common that the stock trading frequency and volume at the end of the day was slightly higher than that of the start of the day, indicating that there was a large amount of stock trading people were trying to get in just before the end of the workday. The middle of the week, particularly Tuesday and Wednesday, depict more microstructure noise (as seen in the more jagged pattern of the line) than the poles of the week, with Fridays portraying the smoothest pattern and lowest microstructure noise (as seen with the smoother pattern of the line). This is most likely due to the more consistent trading activities at the end of the week (with everyone trying to trade before the weekend). The images in Figure 2.2 depict the diurnal patterns for Target (TGT), which follows similar patterns to all other consumer discretionary stocks except for Ford (F), which is noted later in the chapter.

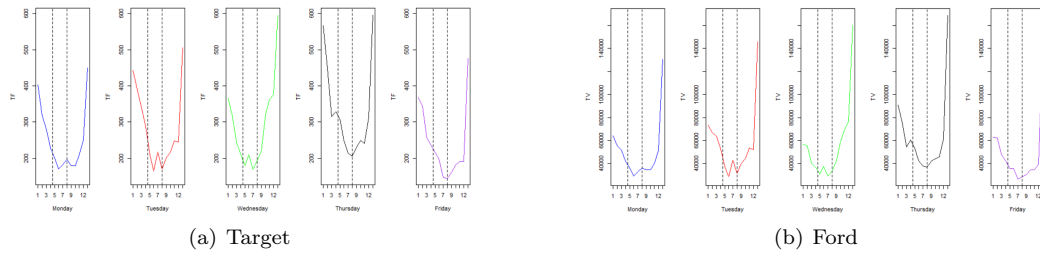


Figure 2.2: Day by Day Time Series for Trading Frequency

All the stocks indicated the same pattern for both trading volume and frequency in the plots of the sample auto-correlations (ACF). For example, see Figure 2.3, which is the the ACF plot for Target for trading volume.

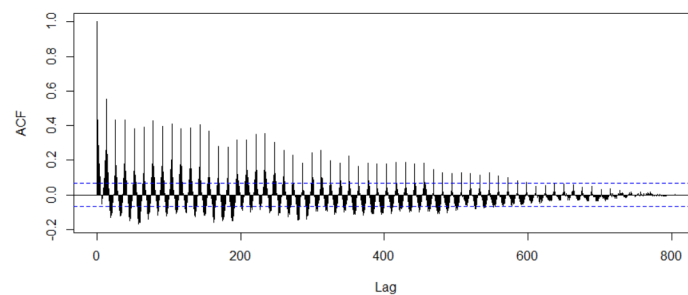


Figure 2.3: Autocorrelation Plot for Target's Trading Volume

There is clearly a periodic relationship between the data and its history, which is present every thirteen data points, or one trading day, for the duration of the data set (all three months of the

fourth quarter). The strength of this relationship decreases over time, but as seen in the graph, the relationship stays significant for a vast majority of the time period in question. This means that past history of trading volume or frequency can be a strong predictor of current and even future trading volume or frequency. We will see this relationship appear again in Chapter 4 when we discuss the bidirectional relationship identified in vector autoregressive models.

2.3.2 Volatility Measures

Realized volatility, range all, and the five minute range, similar to the trading measures, have extreme upper outliers across all stocks. This illustrates the range of actions possible in stock market trading, which causes all 10 stocks to have a skewed right distribution. Also, across all three volatility measures McDonald's has the smallest variance and lowest mean and median. Conversely, across all three measures Amazon has the largest variance and largest mean and median. This indicates that McDonald's is the least volatile stock in the set and Amazon is the most volatile stock in the set. It is interesting to note that although the five minute range is less noisy than range all, both variables have the exact same distributions in each of the 10 stocks. For example, see the distributions of the five minute range for each of the stocks below in Figure 2.4 and Table 2.1.

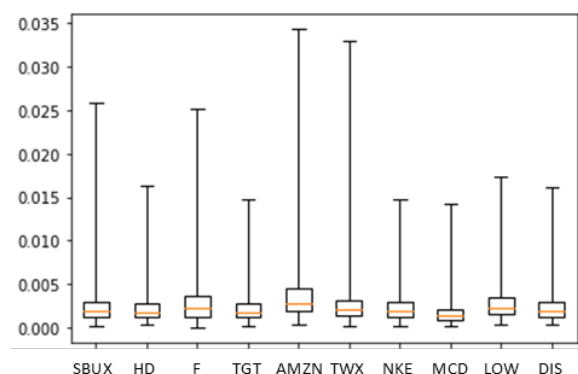


Figure 2.4: Numerical Summaries of Five Minute Range Data for All Stocks

	SBUX	HD	F	TGT	AMZN	TWX	NKE	MCD	LOW	DIS
Min	0.00013	0.000253	0	0.00015	0.000387	0.000142	0.000130	0.000205	0.000398	0.000285
Q1	0.001241	0.001182	0.001207	0.00113	0.001812	0.00132	0.001258	0.000845	0.001473	0.00126
Med	0.001949	0.001769	0.002313	0.00178	0.002748	0.00207	0.001894	0.001317	0.002318	0.00192
Q3	0.002998	0.002840	0.003552	0.0027	0.004433	0.00316	0.002867	0.002012	0.003445	0.00291
Max	0.025776	0.016312	0.02522	0.0148	0.034329	0.033	0.014722	0.014227	0.017343	0.0161

Table 2.1: Numerical Summaries of Five Minute Range Data for All Stocks

For each volatility measure, the plots depict consistent patterns and shapes across all stocks. While the realized volatility had the largest discrepancy between stocks, with about half having one large spike, these measures were significantly more noisy than the frequency and volume. The most consistent pattern across all three volatility measures was that the larger spikes tended to occur in the middle of the week, as evidenced by the plots of Target in Figure 2.5. The most consistency occurred in the range all time series plots, where the shapes and levels of noise across stocks were all relatively the same. Comparatively, realized variance was the least consistent measure.

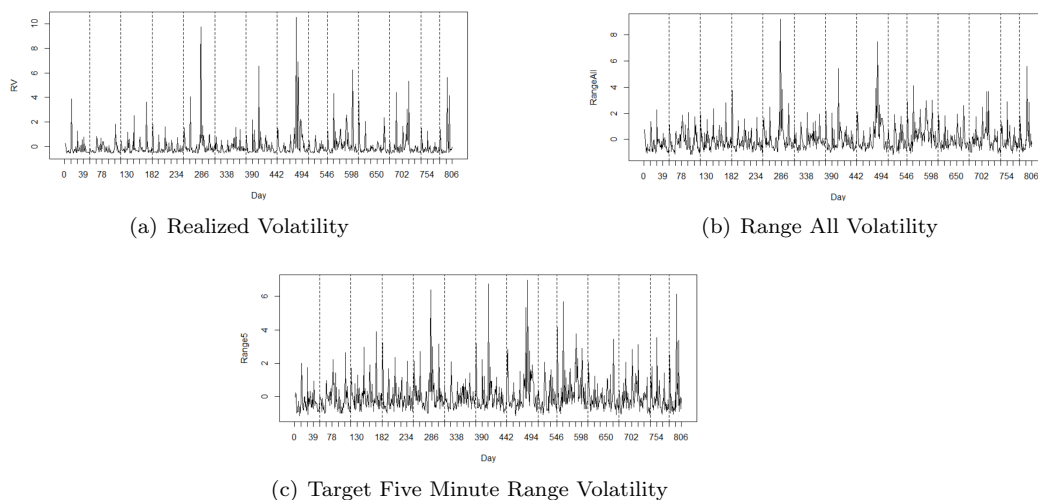


Figure 2.5: Volatility Time Series for Target

The daily time series for our various volatility measures follows a similar formatting to that of trading frequency and volume from the previous section. Each of Target's volatility measures followed a similar daily pattern, making an "L" or backwards "J" shape. This depicts a large variance in volatility in the morning, with a lower, more consistent volatility starting after the conclusion of lunch (as seen following the second dotted line in Figure 2.6). All three measures contained more noise in their patterns in the middle of the week, with Monday's displaying the smoothest pattern. Within the days themselves, the afternoon depicted much more microstructure noise than the early morning, which displayed a fairly smooth negative slope. Just before lunch started (the first dotted line), the patterns became much more jagged. This observation was consistent within the other nine consumer discretionary stocks as well.

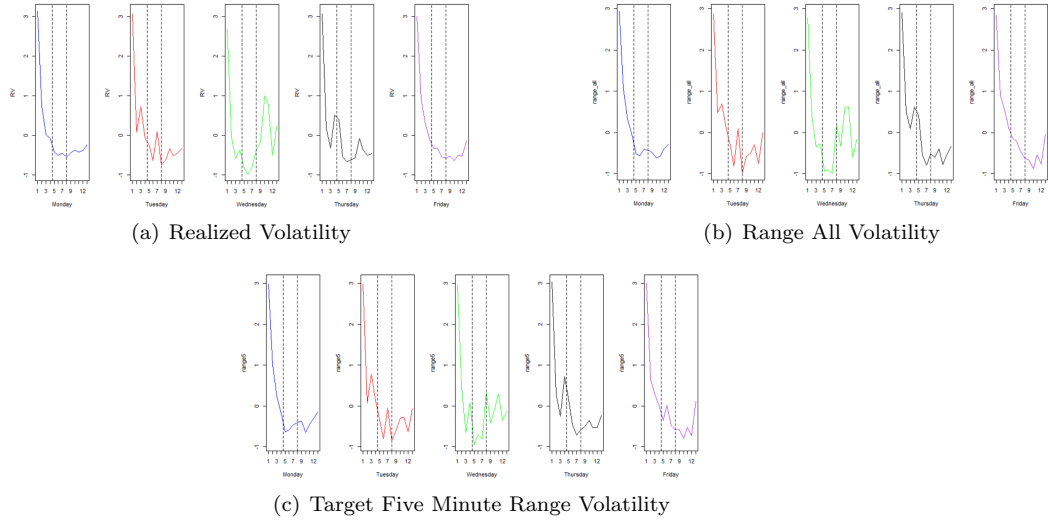


Figure 2.6: Day by Day Time Series for Target

Identically to the ACF plots discussed for the trading variables, all of the stocks indicated the same pattern for all three volatility variables. For example, see Figure 2.7, the ACF plot for Target for the five minute range. This is the same pattern identified for the trading variables, which means that past history of volatility variables can be a strong predictor of current and even future volatility variables, which is one of the main research goals of this project. We will see this relationship appear again in Chapters 3 and 4 when we discuss models to predict volatility using both lagged volatility and trading variables.

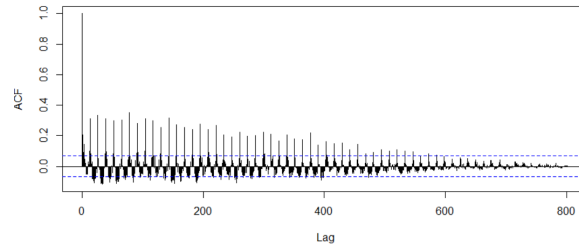


Figure 2.7: Autocorrelation Plot for Target's Five Minute Range Volatility

2.3.3 Relationship between Trading and Volatility Variables

In order to make the most appropriate models, we needed to understand the statistical characteristics of the variables and how they are related with each other over time. Figure 2.8 represents the correlation matrix for the entire cleaned data set. The individual variable distributions are seen across the main diagonal depicting the y-intercept term calculated by R, the volatility measure lagged, prior day rolling average of volatility measure, trading frequency, prior day rolling average of trading frequency, lagged trading frequency, trading volume, prior day rolling average of trading volume, and lagged trading volume as var 1 to var 9, respectively. Each of these variables are explicitly defined in Chapters 3.2 and 4.2. The upper triangle of the matrix depicts the correlation

coefficient between two variables and the lower triangle depicts the scatter plot relationship of the variable pairs.

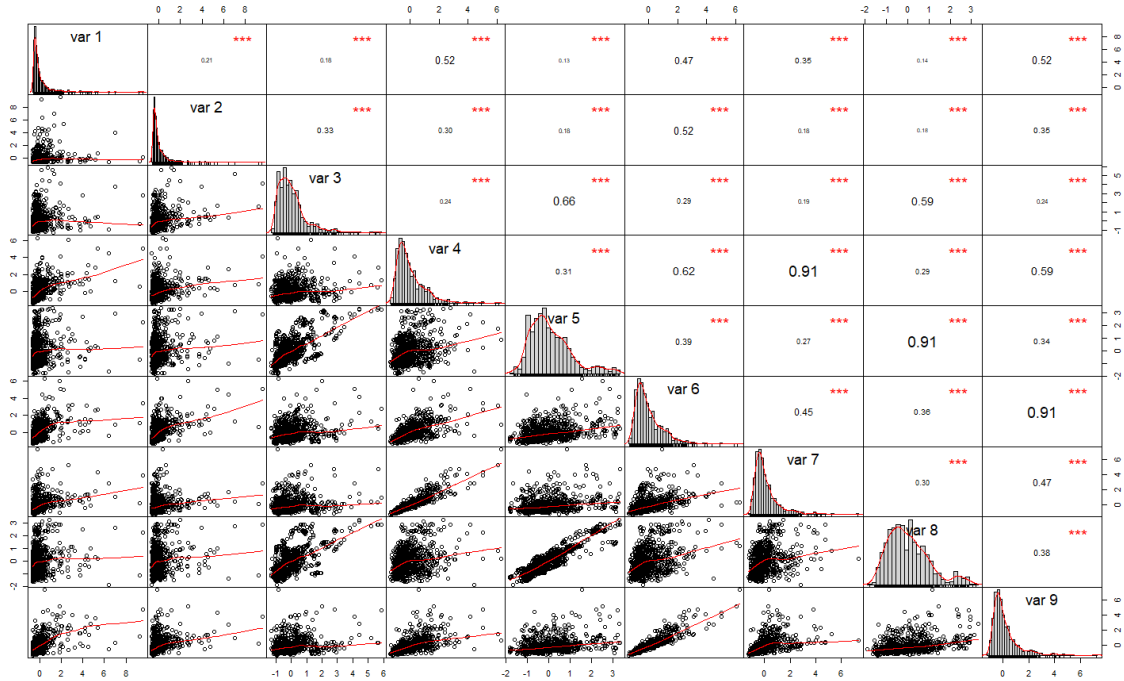


Figure 2.8: Summary Data Correlation Matrix

As explained in previous sections, the data's distributions are each skewed right, which is typical for stock data. All of the correlations amongst the data were significant, but only three correlations were high enough for us to give deliberate attention to avoid using in certain model cases due to their evidence of extreme multicollinearity. This multicollinearity is seen in Figure 2.8 where the correlation coefficients between the trading frequency and trading volume variables were each correlated by 0.91. This relationship is further seen in the scatter plots for these three relationships on the lower triangle, which depict a strong, positive, linear relationship.

Although the trading and volatility variables depict similar diurnal patterns with themselves, when compared to each other we find that their patterns are different. Both trading frequency and volume create a “U” shaped relationship, which mirrors our volatility measures for the morning hours with their high starting values and downward sloping trends. The patterns then become opposite of each other after lunch (around 1:00PM), as seen in Figure 2.9 following the second vertical line. After lunch, trading frequency and volume begin an upward trend whereas the volatility measures plateau or only trend upward slightly. For example, the trading frequency can be compared to realized volatility in Figure 2.9 for their shift in patterns.

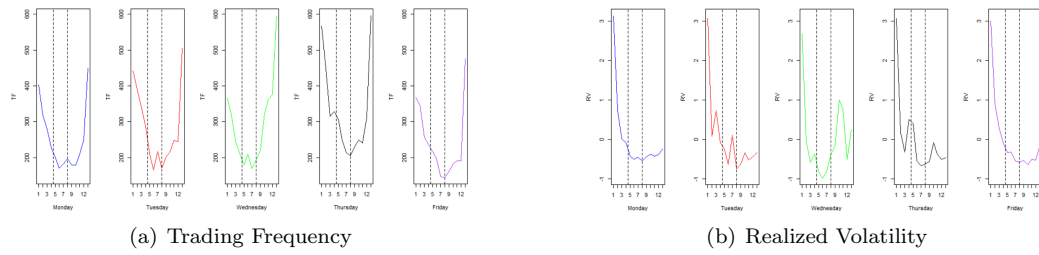


Figure 2.9: Trading Frequency vs. Realized Volatility Day by Day Time Series for Target

Most of our stocks followed this similar pattern aside from Ford. This shift in relationship is also evident in the cross-correlation function (CCF) plots, where the positive correlation exists for the morning increments and shifts to a negative correlation for the afternoon.

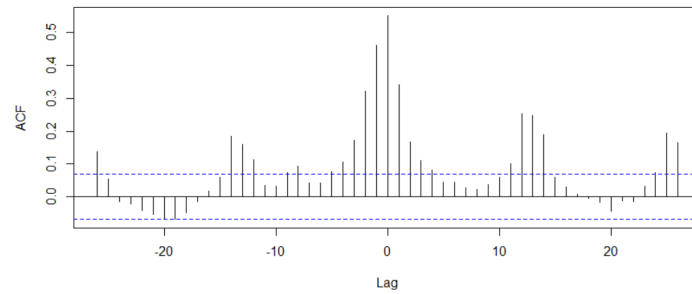


Figure 2.10: Ford CCF Plot Between Trading Volume and Realized Volatility

Interestingly, nine out of ten of our stocks followed this similar relationship between trading and volatility variables with the exception of Ford's stock, depicted in Figure 2.11. Ford's trading frequency follows the same pattern as the volatilities.

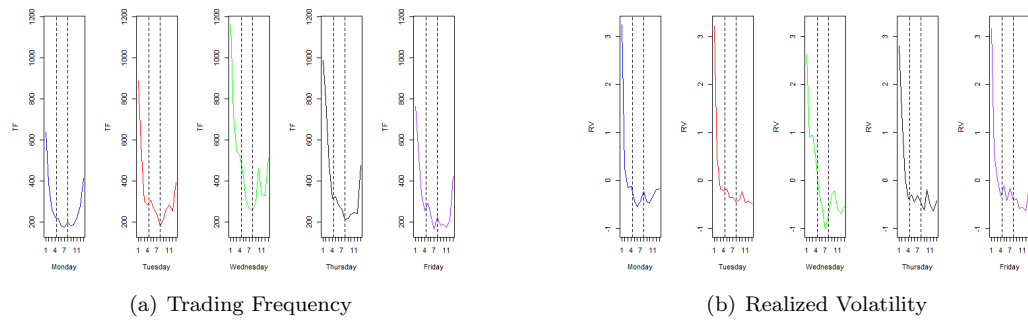


Figure 2.11: Trading Frequency vs. Realized Volatility Day by Day Time Series for Ford

This similarity in pattern is also supported in the CCF plot for Ford, which remains positive, indicating positive correlation and a similar pattern throughout the full workday.

Due to the Ford press conference discussed in Section 2.3.1, on December 18th, 2013, Ford's stock became extremely volatile for several days, which we believe has some effect on the diurnal patterns and relationship between the trading frequency, volume, and its volatilities.

With the exception of Ford, the other nine stocks all show the same relationship between each pair of variables (trading volume and trading frequency each paired with realized volatility, five minute range, and range all) in their CCF plots. For example, see Figure 2.12 for Target between realized volatility and trading volume.

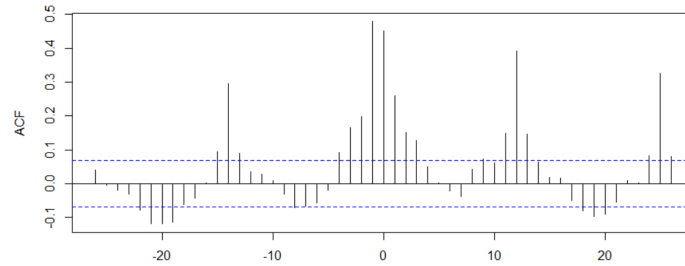


Figure 2.12: CCF plot for Target between trading volume and realized volatility

Chapter 3

Heterogeneous Auto-Regressive Model

In this chapter, we will describe the Heterogeneous Auto-Regressive (HAR) Model: its general form, our covariate implementation, the models we designed, and the results and analysis.

3.1 HAR Overview

The Heterogeneous Auto-Regressive (HAR) Model is a regression model that represents the volatility of a certain stock utilizing the covariates related to trading frequency and volume that we will define below. As discussed in Chapter 1.1, the HAR model takes into account long-memory volatility, which allows a more accurate model of our sporadic financial data (Corsi, 2009). Thus, the HAR model could well describe the dynamic relationship between the volatility measure and the trading frequency and trading volume by combining data of different frequencies, and it allows us to look into the past without having to estimate a large number of lagged terms.

The generic equations described below are inspired by the research of Corsi (2009) for the HAR-RV model and are used to predict V_t , the variable of interest at time t :

$$V_t = c + \beta V_{t-1} + \beta^d V_{t-1}^d + \epsilon_t, \quad (3.1)$$

$$V_t = c + \beta V_{t-1} + \beta^d V_{t-1}^d + \omega \hat{f}_t + \epsilon_t, \quad (3.2)$$

$$V_t = c + \beta V_{t-1} + \beta^d V_{t-1}^d + \omega \hat{f}_t^v + \epsilon_t, \quad (3.3)$$

for $t = 14, 15, \dots, T$, $V_{t-1}^d = 13^{-1} \sum_{l=1}^{13} V_{t-l}$ aggregates the variable over 6.5 hours (i.e. one day). Further, c is a constant, the β , β^d , and ω terms are regression coefficients, $\hat{f}_t$ and $\hat{f}_t^v$ are the trading frequency and volume at time t , respectively, and the ϵ terms are irreducible error. In the following sections we will describe how we implement this concept in our project, with our own version of the general models described in Chapter 3.3 using the covariates described in Chapter 3.2.

3.2 Implementation

To use the HAR model we will forgo the HAR package available in R and instead create our own models using the multiple linear regression function. Before we discuss creation of our models we will first define our covariates. These predictors relate to the volatility measures (which are defined in Chapter 2.2), trading frequency, and trading volume, and are defined as follows:

- **Lagged Volatility:** V_{t-1} generally; RA_{t-1} , $R5_{t-1}$, or RV_{t-1} depending on the volatility response variable.
- **Prior Day Rolling Average Volatility:** V_{t-1}^d generally; RA_{t-1}^d , $R5_{t-1}^d$, or RV_{t-1}^d where $RA_{t-1}^d = 13^{-1} \sum_{l=1}^{13} RA_{t-l}$, for $t = 14, \dots, T$ and R5 and RV follow similar patterns and are used with their respective volatility response variable.
- **Current Trading Volume:** TV_t
- **Lagged Trading Volume:** TV_{t-1}
- **Prior Day Rolling Average Trading Volume:** TV_{t-1}^d where $TV_{t-1}^d = 13^{-1} \sum_{l=1}^{13} TV_{t-l}$, for $t = 14, \dots, T$
- **Current Trading Frequency:** TF_t
- **Lagged Trading Frequency:** TF_{t-1}
- **Prior Day Rolling Average Trading Frequency:** TF_{t-1}^d where $TF_{t-1}^d = 13^{-1} \sum_{l=1}^{13} TF_{t-l}$, for $t = 14, \dots, T$

The relationships between these covariates is shown in Figure 2.8.

The first model we will discuss encompasses all of our possible predictors. The subsequent models are derived from there. While calibrating our models we overcome autocorrelation and heteroskedasticity in the errors by running the Newey-West procedure on each model. This adjusts the standard error of the coefficients and their subsequent significance to avoid the detrimental effects of autocorrelation and heteroskedasticity within a given model (Hanck and Arnold, 2020). We will discuss each of our final models in more detail below.

3.3 Models

We begin by creating and examining a total of 18 models (six models for each of the three volatility measures) that include various combinations of covariates. We will center our focus on these six models because they explain the most information about the relationships between the volatility measures, trading volume, and trading frequency. Below are the general versions of the models we created.

Equation (3.4) refers to our first model, HAR.All. HAR.All contains all of the possible covariates in order to see a full picture of how well the current volatility could be modeled based off of the lagged volatility, the prior day rolling average volatility, the current trading volume, the lagged trading volume, the prior day rolling average trading volume, the current trading frequency, the lagged trading frequency, and the prior day rolling average trading frequency values.

$$V_t = c + \beta V_{t-1} + \beta_2 V_{t-1}^d + \beta_3 TV_t + \beta_4 TV_{t-1} + \beta_5 TV_{t-1}^d + \beta_6 TF_t + \beta_7 TF_{t-1} + \beta_8 TF_{t-1}^d + \epsilon_t. \quad (3.4)$$

Equation (3.5) refers to our second model, HAR_NoRoll, which removes all of the prior day rolling averages from the equation. It includes the following covariates: the lagged volatility, the current trading volume, the lagged trading volume, the current trading frequency, and the lagged trading frequency.

$$V_t = c + \beta V_{t-1} + \beta_2 TV_t + \beta_3 TV_{t-1} + \beta_4 TF_t + \beta_5 TF_{t-1} + \epsilon_t. \quad (3.5)$$

Equation (3.6) refers to our third model, HAR_TV, which includes the lagged volatility, the current trading volume, the lagged trading volume, and the prior day rolling average for trading volume. This model helps depict what happens when trading frequency is completely removed from the equation.

$$V_t = c + \beta V_{t-1} + \beta_2 TV_t + \beta_3 TV_{t-1} + \beta_4 TV_{t-1}^d + \epsilon_t. \quad (3.6)$$

Equation (3.7) refers to our fourth model, HAR_TF, which includes the lagged volatility, the current trading frequency, the lagged trading frequency, and the prior day rolling average for trading frequency. This model helps depict what happens when trading volume is completely removed from the equation.

$$V_t = c + \beta RA_{t-1} + \beta_2 TF_t + \beta_3 TF_{t-1} + \beta_4 TF_{t-1}^d + \epsilon_t. \quad (3.7)$$

Equation (3.8) refers to our fifth model, HAR_TVTF, which includes the current trading volume, the lagged trading volume, the prior day rolling average for trading volume, the current trading frequency, the lagged trading frequency, and the prior day rolling average for trading frequency. This model helps to explain what is depicted when there is no record of the previous volatility in the equation.

$$V_t = c + \beta TV_t + \beta_2 TV_{t-1} + \beta_3 TV_{t-1}^d + \beta_4 TF_t + \beta_5 TF_{t-1} + \beta_6 TF_{t-1}^d + \epsilon_t. \quad (3.8)$$

Equation (3.9) refers to our sixth and final model, HAR_TVTF_noroll, which includes the current trading volume, the lagged trading volume, the current trading frequency, and the lagged trading frequency.

$$V_t = c + \beta TV_t + \beta_2 TV_{t-1} + \beta_3 TF_t + \beta_4 TF_{t-1} + \epsilon_t. \quad (3.9)$$

These six models are computed for each of the three volatility measures: range all (displayed in the models above), five minute range, and realized volatility, resulting in a total of 18 models.

3.4 Results and Analysis

Our HAR models help to illustrate the relationships between the volatility, trading frequency, and trading volume. In Figure 3.1 below, we illustrate the significant coefficients (at various levels) for each of the six models for the range all volatility measure for Target (TGT), our example stock:

TGT RangeAll Model	Lagged Volatility	Prior Day Rolling Average Volatility	Current Trading Frequency	Prior Day Rolling Average Trading Frequency	Lagged Trade Frequency	Current Trade Volume	Prior Day Rolling Average Trading Volume	Lagged Trading Volume	R ²	Adjusted R ²	Residual Standard Error
HAR_all	.	***	***	***	***	***	*	***	0.5261	0.5212	0.6919
HAR_noroll	***		***		***	***		***	0.5056	0.5025	0.7053
HAR_TV			***	**	.				0.4104	0.4074	0.7698
HAR_TF	**					***	*	***	0.3738	0.3706	0.7934
HAR_TVTF			***	***	**	***	**	***	0.5076	0.5039	0.7044
HAR_TVTF_noroll			***		***	***		***	0.4845	0.4819	0.7198

Symbol	α Level
***	0.001
**	0.01
*	0.05
.	0.1

Figure 3.1: HAR Covariate Significance and Model Fit for Target's Range All Volatility

Each column in Figure 3.1 depicts the covariates used and each row is a model. Covariates that are not included as predictors in the selected model are represented by grey boxes, for example, the three grey boxes in the HAR_noroll show that the prior day rolling averages for volatility, trading volume, and trading frequency are excluded from the model. Any instances that are white and blank were included in the model are not significant at any level. As the “Symbol α ” chart explains, the yellow “.” instances are significant at $\alpha = 0.1$, the purple “*” instances are significant at $\alpha = 0.05$, the blue “**” instances are significant at $\alpha = 0.01$, and the green “***” instances are significant at $\alpha = 0.001$.

The numerical results of our HAR model for TGT are as follows:

Responses/Predictors	V_{t-1}	Rolled V_{t-1}	TF	Rolled TF_{t-1}	TF_{t-1}	TV	Rolled TV_{t-1}	TV_{t-1}
HAR_all	0.1066	0.1390	1.1162	-0.4190	-0.4731	-0.5775	0.1979	0.5778
HAR_noroll	0.1889		1.0429	-0.6799		-0.5320	0.7074	
HAR_TF	0.0844		0.5118	-0.1735	0.2072			
HAR_TV	0.1397					0.2683	-0.0971	0.4080
HAR_TVTF			1.1423	-0.4794	-0.3231	-0.6017	0.3474	0.4906
HAR_TVTF_noroll			1.0542		-0.4793	-0.5475		0.6116

Table 3.1: HAR Covariate Coefficients for Target’s Range All Volatility

To facilitate ease of understanding in our results we will discuss Target (TGT) as our main example and point out any notable discrepancies among other stocks. Model HAR_all performs the strongest in terms of adjusted R^2 values. As shown above, model HAR_all is the only model to include every possible predictor, and thus having the most information to predict the volatility measure. After model HAR_all, models HAR_noroll, HAR_TVTF, and HAR_TVTF_noroll also perform well. Between volatility measures, these three models were almost interchangeable, with all having similar adjusted R^2 values and numbers of significant coefficients. Models HAR_TF and HAR_TV are the lowest performing models, with HAR_TV consistently performing worse than HAR_TF across each volatility measure. From this, we can conclude that while including frequency is more beneficial than just volume, in order to produce the strongest results it is important to have a combination of the two trading variables.

Another portion of our analysis observes the significance levels of each variable in the models. For Target, there are two covariates that are significant at the $\alpha = 0.001$ level across every model and volatility measure: trade frequency and lagged trade volume. The current trade volume is significant at the $\alpha = 0.001$ level for every model except HAR_TV for the realized volatility, where it is only significant at the $\alpha = 0.01$ level. The only covariates to hold no significance in any model are the lagged trading frequency and the lagged response variable. Lagged trading frequency is significant for all but HAR_TF for realized volatility and the lagged response variable is only significant for models using the range all response variable. From this we can conclude that every variable, except the lagged response, is contributing to the overall success of the models. We can also conclude that the lagged volume is more significant than the rolled volume, but the rolled frequency is more significant than the lagged frequency.

From looking at the coefficients table, we can see that the trading frequency has the coefficient with the largest impact in every model. All of the coefficients for trading frequency are positive, and all but one are above 1.0. Trading frequency is the only predictor to have any coefficients over 1.0. The rolled trading frequency has a consistently negative correlation with volatility across models, although none of the coefficients is extremely impactful. Between lagged terms, lagged volatility is the least impactful, with values between 0.0844 and 0.1889. Lagged trading volume is more impactful, with values ranging between 0.4080 and 0.6116. For both lagged volatility and lagged trading volume the relationship with current volatility was positive across all models. For lagged trading frequency the values had a larger range, between -0.4793 and 0.2072. The relationship

between lagged trading frequency and current volatility was negative for three out of the four models it was included in. From all of this we can conclude that higher current trading frequency will lead to increased volatility, but higher frequency from the previous 30 minute interval or average over the previous day will lead to lower volatility. We can also conclude higher previous volatility will continue to increase current volatility. The last conclusion we draw is that larger trading volumes lead to lower volatility, but larger trading volumes from the previous 30 minute interval or average over the previous day will lead to higher volatility.

The results from Target are mostly consistent with the results seen across all 10 stocks. For the generalized models that incorporate the summary data from all 10 stocks, all of the observations and conclusions drawn above remain. The same is true for McDonald's, which is the least volatile stock.

However, Amazon, the most volatile stock, has some important distinctions in its model results. Although the best performing models (via the adjusted R^2 values) remain the same, the margins are much slimmer, indicating that high volatility may reduce the effectiveness of one predictor over another. This is seen via the significant predictors as well, notably being that no predictor is consistently significant across the different models and volatility measures.

Similarly, there are interesting observations for several other stocks that deviate from the consistent results. Ford, for example, has relatively few significant predictors. However, trading frequency and trading volume are both significant in most cases. Starbucks also has trading volume as a consistent significant predictor, along with the lagged trading frequency, and relatively few significant predictors overall. Finally, Time Warner is the only stock where the lagged response variable (volatility measure) is consistently significant.

Chapter 4

Vector Auto-Regressive Model

In this chapter, we will discuss the use of the Vector Auto-Regressive (VAR) Model in studying the cross dependence between intraday volatility and trading volume and trading frequency over time.

4.1 VAR Overview

In addition to implementing heterogeneous auto-regressive models, we will also explore the implementation of vector auto-regressive (VAR) models. Lu and Lin (2010) used a bivariate VAR model in their research to further explore the relationship between intraday volatility and trading volume and examine the bidirectionality, or the feedback effect of the relationship. This is motivated from the hypothesis of bidirectionality that was identified in Chapter 2.3. We also have the ability to set p , or the duration of the lag effect that is captured by the model. Lu and Lin (2010) found the model with order 2 to be optimal, or a VAR(2) model. The underlying VAR(p) model (Pfaff and Stigler, 2018) for a k -dimensional vector time series y_t is represented by:

$$y_t = A_1 y_{t-1} + \dots + A_p y_{t-p} + CD_t + u_t. \quad (4.1)$$

Equation (4.1) is the vector form of k multiple linear regression models that are then solved via the ordinary least squares method. Here, the matrices $A_1, \dots, A_p$ are of size $k \times k$ and represent the linear regression coefficients. The term CD_t is an optional $k \times 1$ term that can add either a constant or a seasonal trend to the model. Finally, u_t is a $k \times 1$ error term that captures spherical disturbance.

4.2 Implementation

For this project, we incorporate one lagged term and disregard any manually added terms. This means we set $p = 1$ and $CD_t = 0$, which leaves us with the following VAR(1) model:

$$y_t = A y_{t-1} + u_t. \quad (4.2)$$

For example, take the model VAR_TF, where we examine the relationship between trading frequency and the volatility measures. When we consider this model for Target with the volatility measure of range all, we find $k = 3$ since we have three variables (range all, trading frequency, and prior day rolling average trading frequency). The equation above is then solved using the ordinary least

squares on each row of the matrix using the data set provided (all three months of trading days, i.e., t ranges from 2 to 793). The matrix A for this model can be seen in Chapter 4.4. Further, we are able to examine the significance of each of these coefficients and draw conclusions regarding the relationships between the variables, as seen in Chapter 4.4.

We will implement two VAR models using the VAR package in R, which follows the generalized model discussed above. We will use models that minimize the effects of collinearity between independent variables by only including predictors that do not have a high degree of correlation between them, as seen in the correlation matrix in Figure 2.8. For example, trading frequency and trading volume are strongly correlated and thus are not considered in the same model together. The covariates present in these models are defined below.

- **Volatility:** V_t generally; RA_t , $R5_t$, or RV_t depending on the volatility response variable.
- **Current Trading Volume:** TV_t
- **Prior Day Rolling Average Trading Volume:** TV_{t-1}^d where $TV_{t-1}^d = 13^{-1} \sum_{l=1}^{13} TV_{t-l}$, for $t = 14, \dots, T$
- **Current Trading Frequency:** TF_t
- **Prior Day Rolling Average Trading Frequency:** TF_{t-1}^d where $TF_{t-1}^d = 13^{-1} \sum_{l=1}^{13} TF_{t-l}$, for $t = 14, \dots, T$

4.3 Models

For the VAR(1) models, we set up y_t to include the volatility measure, the current value of the trading measure, and the prior day rolling average of the trading measure. That is, our first model, VAR-TV, included the following three variables:

$$\text{Model VAR-TV: } y_t = \begin{bmatrix} V_t \\ TV_t \\ TV_{t-1}^d \end{bmatrix}. \quad (4.3)$$

Our second model, VAR-TF, is similar to our first except it uses trading frequency instead of trading volume, resulting in the following y_t :

$$\text{Model VAR-TF: } y_t = \begin{bmatrix} V_t \\ TF_t \\ TF_{t-1}^d \end{bmatrix}. \quad (4.4)$$

These two models are investigated across each of the three volatility measures resulting in six different VAR models for each stock.

4.4 Results and Analysis

Our VAR models help to illustrate the potential bidirectional relationships between the volatility, trading frequency, and trading volume. Below, in Figure 4.1, the results are illustrated for each of the models for the range all volatility measure for Target (TGT), our example stock. We also see the coefficients for the VAR-TF model and the VAR-TV model in their respective coefficient matrices below, which are also important when drawing conclusions.

Model	Response Predictors	Current Volatility (RA)	Current Trading Volume	Prior Day Rolling Average Trading Volume	R ²	Adjusted R ²	Residual Standard Error
VAR_TV	Current Volatility (RA)	***	***		0.3163	0.3137	0.828
	Current Trading Volume		***	***	0.2174	0.2144	0.8869
	Prior Day Rolling Average Trading Volume	***	***	***	0.9835	0.9834	0.1288
VAR_TF	Current Volatility (RA)	*	***	*	0.2596	0.2568	0.8616
	Current Trading Frequency		***	***	0.4326	0.4305	0.7552
	Prior Day Rolling Average Trading Frequency	.	***	***	0.9885	0.9885	0.1074

Symbol	α Level
***	0.001
**	0.01
*	0.05
.	0.1

Figure 4.1: VAR Covariate Significance and Model Fit for Target's Range All Volatility

Figure 4.1 depicts the significance of each predictor at the various α 's we ran in our models. Each of the columns depicts the covariates used, in this case they are all lagged by one time period due to the nature of the VAR model. The chart is also divided into two models: **VAR_TV** and **VAR_TF**, with the rows within depicting the $k = 3$ equations VAR creates. Any instances that are white are not significant at any level in the model. As the "Symbol α " chart explains, the yellow "." instances are significant at $\alpha = 0.1$, the purple "*" instances are significant at $\alpha = 0.05$, the blue "**" instances are significant at $\alpha = 0.01$, and the green "***" instances are significant at $\alpha = 0.001$.

The estimated coefficient matrix A for the model **VAR_TF** is as follows, as described in Chapter 4.2:

Responses/Predictors	range all _{$t-1$}	TF _{$t-1$}	rolled TF _{$t-1$}
range all _{t}	0.0881	0.4784	-0.0727
TF _{t}	0.0042	0.5804	0.1435
rolled TF _{t}	0.0090	0.0916	0.9494

Table 4.1: Matrix A for Model VAR_TF

Similarly, the estimated coefficient matrix A for the model **VAR_TV** is as follows:

Responses/Predictors	range all _{$t-1$}	TV _{$t-1$}	rolled TV _{$t-1$}
range all _{t}	0.1501	0.4872	-0.0231
TV _{t}	0.0522	0.3482	0.1796
rolled TV _{t}	0.0315	0.0755	0.9573

Table 4.2: Matrix A for Model VAR_TV

4.4.1 Predicting Volatility Measures

First, we will examine the prediction of the volatility measures using VAR(1). This is the prediction of the first row of y_t , which is V_t , the volatility measure. For Target, as seen in Figure 4.1, the coefficients of the stock's current frequency and volume terms are significant at the $\alpha = 0.001$ level across all six models. This is supported by observing the coefficients of each of these terms in

the above coefficient matrices, where the current frequency and volume terms have coefficients with significantly greater absolute value than the other predictors. Additionally, across all three volatility measures, model [VAR_TV](#) always performs better than model [VAR_TF](#) based on adjusted R^2 values, meaning that volume is more useful than frequency in predicting volatility. A distinction between models is that the prior day rolling average trading volume is always significant for model [VAR_TV](#), but the prior day rolling average trading frequency is never significant for model [VAR_TF](#). This indicates that the prior day rolling average trading volume is useful in predicting volatility but the prior day rolling average trading frequency is not.

These conclusions are consistent across most of the stocks that were analyzed. In particular, all three conclusions above hold for the models based on the summary data from all ten stocks. Further, both Amazon (most volatile) and McDonald's (least volatile) have their current frequency or volume term remain significant (at $\alpha = 0.001$) across both models and all three volatility measures. Conversely, although the [VAR_TV](#) model outperforms the [VAR_TF](#) model for McDonald's (similar to the majority of stocks), [VAR_TF](#) performs better for Amazon (based on adjusted R^2 values). Finally, the significance of the prior day rolling average trading frequency or volume term varies for both Amazon and McDonald's. While it is never significant for Amazon, it is only significant for realized volatility for McDonald's. Overall, these results do not indicate any striking trends between significance of the trading variables and the volatility of the stock in question.

4.4.2 Predicting Trading Measures

Second, we will examine the prediction of the current trading measures and prior day rolling average trading measures using VAR. This is the prediction of the second and third rows of y_t , which are TV_t and TV_{t-1}^d or TF_t and TF_{t-1}^d , depending on the model. This is the feedback effect of using the lagged terms in the model to predict the current values of the predictors. For Target, across all three volatility measures in model [VAR_TV](#), the models do a poor job of predicting the current trading volume relative to predicting the current prior day rolling average trading volume, with the adjusted R^2 values falling between 0.21 and 0.22 relative to the adjusted R^2 values of approximately 0.98.

Similarly, across all three volatility measures for Target's model [VAR_TF](#), the models are better at predicting the current prior day rolling average trading frequency than predicting the current trading frequency, with the adjusted R^2 values of approximately 0.99 relative to values of approximately 0.43. However, the model fit of [VAR_TF](#) for the prediction of current trading frequency has a higher adjusted R^2 value compared to the prediction of the current trading volume in model [VAR_TV](#) (approximately 0.43 versus approximately 0.21). This tells us that the models are strongly bidirectional with respect to the current day rolling average trading measure, and moderately bidirectional with respect to the current trading measure.

These conclusions are consistent with the patterns evident in the other stocks we analyzed, notably the summary data from all 10 stocks, McDonald's, and Amazon. Although Amazon is consistent with these observations, we do see that based on the adjusted R^2 values, there is no notable difference between the two VAR models when predicting the current trading measures.

Chapter 5

Conclusions

This project aimed to explore the relationship between trading volume, trading frequency, and volatility at the intraday level. Ultimately, we focused on analyzing this relationship in the fourth quarter of 2013 on the ten stocks in the consumer discretionary sector of the S&P 100. Before we revisit the research questions identified in Chapter 1.2, it is important to recall that measuring volatility is not strictly defined. To approximate the true volatility of the market using a volatility measure, separate research is required. Instead, we created three measures of volatility to use while exploring this relationship: range all, five minute range, and realized volatility, and explored the aforementioned relationship for each measure.

Using the detailed results for both the heterogeneous autoregressive models and vector autoregressive models, as described in Chapters 3 and 4, respectively, we are able to discuss how these results pertain to our research questions.

1. What is the dynamic relationship between trading volume, trading frequency, and volatility at the intraday level?

In the HAR models, we see that trading frequency is a more helpful predictor of all three volatility measures than trading volume, indicating that frequency may contain more information about the volatility measures than the volume. Further, we saw that the lagged and prior day rolling average terms that were included across the set of models improved performance, indicating that these predictors are also helpful when predicting volatility and do contain more information than just including a trading volume or trading frequency term.

In the VAR models, both trading frequency and trading volume were important and helpful when predicting all three volatility measures. Here, the trading volume was more helpful than the trading frequency at predicting the concurrent volatility measure, and the prior day rolling average of trading volume was usually significant but the corresponding frequency term was not. This indicates that trading volume contains more information than trading frequency when predicting the volatility measures.

The VAR models also indicates the presence of a bidirectional relationship. The models are able to predict the prior day rolling average trading measure with an extremely high level of correlation (with the adjusted R^2 values > 0.95). The models also significantly predict the current trading measures, albeit with a lower correlation.

Overall, in both the HAR models and VAR models, we examined the impact of the stock's overall volatility on the conclusions described above by comparing the results to the findings

of the least volatile stock, McDonald's, and the most volatile stock, Amazon. Interestingly, there was no obvious trends or influence on the overall volatility of the stock impacting model performance or significance.

VAR and HAR models also confirmed that both trading volume and trading frequency are related to the volatility measures and can be described using linear regression models that incorporate the trading measures and their related lagged and prior day rolling average terms. This supports the conclusions we see in the research by Lu and Lin (2010), Cheriyan and Lazar (2019), and Chevallier and Sevi (2012).

2. What models can effectively summarize and explain this relationship?

On the summary data, several HAR models produced an adjusted R^2 value over 0.7 and all six VAR models produced an adjusted R^2 value over 0.5. This illustrates the pattern seen throughout all of the results that both HAR and VAR models can effectively summarize and explain the relationship between trading volume, trading frequency, and the volatility measures.

Further, the inclusion of the lagged terms in both models (manually in HAR and by design in VAR) indicates that this relationship can be modeled using linear regression models that incorporate lagged information about the predictors. This confirms the trends seen in the ACF plots in Chapter 2.3 that indicates this relationship may be present as well as the implementation of VAR(2) by Lu and Lin (2010) and the implementation of HAR-RV by Corsi (2009) in their works.

Finally, it is interesting to note a distinction between the sets of models. We discussed in the answer to the previous research question that trading frequency appeared to contain more information for the HAR models and trading volume appeared to contain more information for the VAR models. We would expect that this would be consistent across the two model types and thus more research is required to investigate this discrepancy.

Ultimately, we confirmed that the trading volume, trading frequency, and volatility relationship exists and can be modeled by time series methods. However, our research also poses more questions that require further exploration. Hopefully this inspires future research into this dynamic relationship to help describe, understand, and predict volatility in the stock market.

Chapter 6

Next Steps

After concluding our research for this project, we have created a list of potential next steps that we recommend a future research team or project team complete. These next steps look at both longer time spans to further test our conclusions and alternative model designs to potentially better encompass the relationship between trading volume, trading frequency, and volatility in the stock market.

6.1 Alternative Data Cleaning Methods

If further research is completed, we would recommend revisiting the data cleaning methodology and spend more time making it adaptable to different queries. This recommendation is inspired by a conversation we had with James Kingsley, a Computational Scientist from the Academic and Research Computing group (ARC) at WPI. In our discussions, we compared the different speeds and preferable qualities of several different data cleaning platforms. Ultimately, the strongest candidate was to create a SQL database that would support repetitive querying and more efficiently create covariates. Unfortunately, due to time constraints we were not able to make these adjustments and shift from our Python cleaning code, so we recommend another project team makes these preferred adjustments before continuing research.

6.2 Expanding the Time Frame

Due to time and resource constraints, we only examined the intraday patterns for the fourth quarter of 2013. We recommend additional research be done by expanding the timeline to the entire year of 2013, or even longer. With this, more data points would be available for examining the intraday trends, and it opens the possibility of doing day-to-day trend analysis as well. With this, some potential research questions include:

- Do specific holidays impact the volatility of the market?
- How significant is the impact of major company benchmarks or current events on the stocks' volatility?
- Is there a seasonal effect on the volatility of the market?

6.3 Dividing the Model

After doing further analysis on the diurnal patterns found in Chapter 2.3, we noticed the positive correlation between the trading variables and volatility for the morning hours (before the end of lunch). Interestingly, this correlation becomes negative (with the trading variables trending positive and the volatility variables plateauing) after lunch ends. Because of this shift in the relationship, we were curious if separate models for the morning and afternoon trades would better model the relationship.

We ran a miniature simulation to test this theory using the Lowe's stock with a very simple model regressing volatility on the current trading frequency and current trading volume, but divided the model up into the morning 30 minute intervals and the afternoon 30 minute intervals. The morning observations consisted of seven out of the 13 daily observations from 9:30AM to 1:00PM and the afternoon observations encompassed the remaining six daily observations, from 1:30PM to 4:00PM. The model can be seen below.

$$V_t = c + \beta_1 TV_t + \beta_2 TF_t + \epsilon_t. \quad (6.1)$$

The results were inconclusive, as our new models were a better fit for the morning, but was much weaker for the afternoon, as can be seen in the table below.

Model	β_1	β_2	R^2	Adjusted R^2	RSE
Entire Day	-0.0060	0.0106	0.404	0.3997	0.7748
Morning Only	-0.0038	0.0117	0.4667	0.4605	0.7345
Afternoon Only	-0.0155	0.0108	0.1856	0.1704	0.9108

Table 6.1: Morning and Afternoon Split Model Results

The morning only model shows improvements to the model fit, with an increased adjusted R^2 and decreased error when compared to the entire day model. In contrast, however, the afternoon only model depicts a weaker model fit and more error. We recommend further research be done to examine if significant improvement for the model can be done by observing the morning and afternoons separately, potentially with one of our models that performed better or included other predictors.

6.4 Examining a Change in TF/TV Covariate

Due to the nature of volatility, people are likely to buy and sell more shares when they either fear the market or are confident that its prospects are good. This is often based on the trends (or change over time) of the trading frequencies or volumes. Because of this, we recommend further research be done using a ΔTF or ΔTV as a covariate, where the Δ of the trading measure indicates the change in value of the measure over a given period of time. This will help to see if, for example, a larger ΔTF indicates more volatility or change in the market trends. This would directly incorporate the effect of change over time of the TV or TF into the model.

6.5 Forecasting

We recommend a future group look into forecasting volatility based on the trading frequencies and volume, potentially adding in some of our other covariates from the prior recommendations.

Bibliography

- Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2009). Realized kernels in practice: Trades and quotes. *The Econometrics Journal*, 12:C1–C32.
- Cheriyian, N. K. and Lazar, D. (2019). Relationship between liquidity, volatility and trading activity: An intraday analysis of indian stock market. *International Journal of Economics and Financial Issues*, 9(1):17.
- Chevallier, J. and Sévi, B. (2012). On the volatility–volume relationship in energy futures markets using intraday data. *Energy Economics*, 34(6):1896–1909.
- Corporate Finance Institute (2021). New York Stock Exchange (NYSE) - Overview, History, Trading. <https://corporatefinanceinstitute.com/resources/knowledge/trading-investing/new-york-stock-exchange-nyse/>. CFI Education Inc.
- Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, 7(2):174–196.
- Fidelity (2021). Consumer discretionary. https://eresearch.fidelity.com/eresearch/goto/markets_sectors/sectors/sectors_in_market.jhtml?tab=learn§or=25#:~:text=The%20Consumer%20Discretionary%20Sector%20encompasses,textiles%20%26%20apparel%20and%20leisure%20equipment. FMR LLC.
- Hanck, C. and Arnold, M. (2020). *Introduction to Econometrics with R*. University of Duisburg-Essen. <https://www.econometrics-with-r.org/15-4-hac-standard-errors.html>.
- Isidore, C. (2013). New Ford models to hit profits, shares fall. <https://money.cnn.com/2013/12/18/news/companies/ford-profit/index.html>. Cable News Network.
- Lu, W.-C. and Lin, F.-J. (2010). An empirical study of volatility and trading volume dynamics using high-frequency data. *The International Journal of Business and Finance Research*, 4(3):93–101.
- Mitchell, C. (2020). Common intra-day stock market patterns. <https://www.thebalance.com/common-intra-day-stock-market-patterns-1031456#:~:text=11:45a.m.\0T1\textendash1:,rangeestablishedduringlunchhour.> The Balance.
- Pfaff, B. and Stigler, M. (2018). Package ‘vars’. <https://cran.r-project.org/web/packages/vars/vars.pdf>. The Comprehensive R Archive Network.
- S&P Dow Jones Indices (2021). S&P 100. <https://www.spglobal.com/spdji/en/indices/equity/sp-100/#overview>. SP 100 - S&P Dow Jones Indices.

Appendix A

S&P 100 Stocks

A.1 Full List of Stocks in the S&P 100

There are currently 101 stocks in the S&P 100 since Google (Alphabet Inc.) is listed twice, once as a class A company and once as a class C company.

Ticker Symbol	Company Name
AAPL	Apple Inc.
ABBV	AbbVie Inc.
ABT	Abbott Laboratories
ACN	Accenture
ADBE	Adobe Inc.
AIG	American International Group
ALL	Allstate
AMGN	Amgen Inc.
AMT	American Tower
AMZN	Amazon.com
AXP	American Express
BA	Boeing Co.
BAC	Bank of America Corp
BIIB	Biogen
BK	The Bank of New York Mellon
BKNG	Booking Holdings
BLK	BlackRock Inc
BMJ	Bristol-Myers Squibb
BRK.B	Berkshire Hathaway
C	Citigroup Inc
CAT	Caterpillar Inc.
CHTR	Charter Communications
CL	Colgate-Palmolive
CMCSA	Comcast Corp.

Ticker Symbol	Company Name
COF	Capital One Financial Corp.
COP	ConocoPhillips
COST	Costco Wholesale Corp.
CRM	salesforce.com
CSCO	Cisco Systems
CVS	CVS Health
CVX	Chevron Corporation
DD	DuPont de Nemours Inc
DHR	Danaher Corporation
DIS	The Walt Disney Company
DOW	Dow Inc.
DUK	Duke Energy
EMR	Emerson Electric Co.
EXC	Exelon
F	Ford Motor Company
FB	Facebook, Inc.
FDX	FedEx
GD	General Dynamics
GE	General Electric
GILD	Gilead Sciences
GM	General Motors
GOOG	Alphabet Inc. (Class C)
GOOGL	Alphabet Inc. (Class A)
GS	Goldman Sachs
HD	Home Depot
HON	Honeywell
IBM	International Business Machines
INTC	Intel Corp.
JNJ	Johnson & Johnson
JPM	JPMorgan Chase & Co.
KHC	Kraft Heinz
KMI	Kinder Morgan
KO	The Coca-Cola Company
LLY	Eli Lilly and Company
LMT	Lockheed Martin
LOW	Lowe's
MA	MasterCard Inc
MCD	McDonald's Corp
MDLZ	Mondelez International
MDT	Medtronic plc
MET	MetLife Inc.
MMM	3M Company
MO	Altria Group
MRK	Merck & Co.
MS	Morgan Stanley
MSFT	Microsoft
NEE	NextEra Energy
NFLX	Netflix
NKE	Nike, Inc.
NVDA	NVIDIA Corp.
ORCL	Oracle Corporation
OXY	Occidental Petroleum Corp.

Ticker Symbol	Company Name
PEP	PepsiCo
PFE	Pfizer Inc
PG	Procter & Gamble Co
PM	Philip Morris International
PYPL	PayPal Holdings
QCOM	Qualcomm Inc.
RTN	Raytheon Technologies
SBUX	Starbucks Corp.
SLB	Schlumberger
SO	Southern Company
SPG	Simon Property Group, Inc.
T	AT&T Inc
TGT	Target Corporation
TMO	Thermo Fisher Scientific
TXN	Texas Instruments
UNH	UnitedHealth Group
UNP	Union Pacific Corporation
UPS	United Parcel Service
USB	U.S. Bancorp
V	Visa Inc.
VZ	Verizon Communications
WBA	Walgreens Boots Alliance
WFC	Wells Fargo
WMT	Walmart
XOM	Exxon Mobil Corp.

A.2 Consumer Discretionary Stocks

The following stocks are in the consumer discretionary sector of the S&P 100 and were used for analysis in this project. Note that references to these stocks throughout the report use the abbreviated company name located in the third column of the table. Also note that Time Warner is not included in the full list of S&P 100 stocks as it was acquired by AT&T in 2018. In 2013, the time period for which we completed the analysis, Time Warner was a distinct company in the S&P 100.

Ticker Symbol	Company Name	Abbreviated Name
AMZN	Amazon.com	Amazon
DIS	The Walt Disney Company	Disney
F	Ford Motor Company	Ford
HD	Home Depot	Home Depot
LOW	Lowe's	Lowe's
MCD	McDonald's Corp	McDonald's
NKE	Nike, Inc.	Nike
SBUX	Starbucks Corp.	Starbucks
TGT	Target Corporation	Target
TWX	Time Warner Inc.	Time Warner

Appendix B

Full HAR Results

The following images depict the covariate significance for each of our eighteen models for each of the ten stocks (plus the summary data). Each stock has three tables, the results for the range all, range 5, and realized volatility volatility measures.

Summary

Summary Range All Model	Lagged Volatility	Prior Day Rolling Average Volatility	Current Trading Frequency	Prior Day Rolling Average Trading Frequency	Lagged Trade Frequency	Current Trade Volume	Prior Day Rolling Average Trading Volume	Lagged Trading Volume	Adjusted R ²	Residual Standard Error
HAR_all		***	***	***	***	***	***	***	0.7208	0.7235
HAR_vol	**		***	***	**	***		***	0.6821	0.6802
HAR_tf			***	***	***	***		***	0.5885	0.5878
HAR_tv	**		***	***		***	***	***	0.6157	0.6150
HAR_tvtf			***	***		***	***	***	0.7063	0.704
HAR_tvtf_vol			***	***	*	***	***	***	0.6814	0.6785

Symbol α Level
 *** 0.001
 ** 0.01
 * 0.05
 . 0.1

Summary Range 5 Model	Lagged Volatility	Prior Day Rolling Average Volatility	Current Trading Frequency	Prior Day Rolling Average Trading Frequency	Lagged Trade Frequency	Current Trade Volume	Prior Day Rolling Average Trading Volume	Lagged Trading Volume	Adjusted R ²	Residual Standard Error
HAR_all		***	***	***	***	***	***	***	0.6767	0.6734
HAR_vol			***	***	***	***		***	0.6484	0.6462
HAR_tf	***			***	***	***		***	0.5648	0.5626
HAR_tv				***		**	***	***	0.6046	0.6026
HAR_tvtf			***	***	*	***	*	***	0.6832	0.6807
HAR_tvtf_vol			***	***	***	***	*	***	0.5474	0.5455

Symbol α Level
 *** 0.001
 ** 0.01
 * 0.05
 . 0.1

Summary RV Model	Lagged Volatility	Prior Day Rolling Average Volatility	Current Trading Frequency	Prior Day Rolling Average Trading Frequency	Lagged Trade Frequency	Current Trade Volume	Prior Day Rolling Average Trading Volume	Lagged Trading Volume	Adjusted R ²	Residual Standard Error
HAR_all	***	***	***	***	***	***	***	***	0.7114	0.7085
HAR_vol			***	***	***	***		***	0.6739	0.6730
HAR_tf	***			***	***	***		***	0.5728	0.5714
HAR_tv				***		*	***	***	0.5842	0.5821
HAR_tvtf			***	***	***	***		***	0.6840	0.6825
HAR_tvtf_vol			***	***	***	***		***	0.6769	0.6742

Symbol α Level
 *** 0.001
 ** 0.01
 * 0.05
 . 0.1

Amazon.com (AMZN)

AMZN Range All Model	Lagged Volatility	Prior Day Rolling Average Volatility	Current Trading Frequency	Prior Day Rolling Average Trading Frequency	Lagged Trade Frequency	Current Trade Volume	Prior Day Rolling Average Trading Volume	Lagged Trading Volume	Adjusted R ²	Residual Standard Error
HAR_all	*	***		***	**	***	***	***	0.545	0.5444
HAR_vol	*			***	***	***		***	0.5277	0.5247
HAR_tf	**		***	*		***		***	0.514	0.5115
HAR_tv	*			***		***		***	0.5245	0.5221
HAR_tvtf				***		*	***	***	0.5329	0.5293
HAR_tvtf_vol				***		***		***	0.5174	0.5149

Symbol α Level
 *** 0.001
 ** 0.01
 * 0.05
 . 0.1

AMZN Range 5 Model	Lagged Volatility	Prior Day Rolling Average Volatility	Current Trading Frequency	Prior Day Rolling Average Trading Frequency	Lagged Trade Frequency	Current Trade Volume	Prior Day Rolling Average Trading Volume	Lagged Trading Volume	Adjusted R ²	Residual Standard Error
HAR_all		***		***	***	***	***	***	0.4605	0.4576
HAR_vol			***	*		***	*	***	0.4507	0.4572
HAR_tf	*			***		***		***	0.4448	0.4419
HAR_tv				***		***		***	0.4575	0.4547
HAR_tvtf				***	**	***	***	***	0.471	0.4669
HAR_tvtf_vol				***	*	***		***	0.4588	0.4542

Symbol α Level
 *** 0.001
 ** 0.01
 * 0.05
 . 0.1

AMZN RV Model	Lagged Volatility	Prior Day Rolling Average Volatility	Current Trading Frequency	Prior Day Rolling Average Trading Frequency	Lagged Trade Frequency	Current Trade Volume	Prior Day Rolling Average Trading Volume	Lagged Trading Volume	Adjusted R ²	Residual Standard Error
HAR_all			*	***	***	***	***	***	0.585	0.5805
HAR_vol			***	**	***	***		***	0.5777	0.575
HAR_tf			***	***	***	***	*	***	0.5312	0.5288
HAR_tv				***	**	***		***	0.5748	0.5727
HAR_tvtf				***	*	***		***	0.5808	0.5778
HAR_tvtf_vol				***	*	***		***	0.4588	0.4542

Symbol α Level
 *** 0.001
 ** 0.01
 * 0.05
 . 0.1

The Walt Disney Company (DIS)

DIS Range4 Model	Lagged Volatility	Prior Day Rolling Average Volatility	Current Trading Frequency	Prior Day Rolling Average Trading Frequency	Lagged Trade Frequency	Current Trade Volume	Prior Day Rolling Average Trading Volume	Lagged Trading Volume	R ²	Adjusted R ²	Residual Standard Error
HAR_M		**	***	**	***	***	***	***	0.4419	0.4382	0.7509
HAR_norol			***	***	***	***	***	***	0.4295	0.4258	0.7577
HAR_TP			***	***	***	***	***	***	0.3156	0.3161	0.5566
HAR_TV			***	***	***	***	***	***	0.2965	0.2929	0.5409
HAR_TVTF			***	***	***	***	***	***	0.4315	0.4272	0.7568
HAR_TVTF_norol			***	***	***	***	***	***	0.4283	0.4264	0.7593

Symbol α Level

*** 0.001

** 0.01

* 0.05

0.1

DIS Range5 Model	Lagged Volatility	Prior Day Rolling Average Volatility	Current Trading Frequency	Prior Day Rolling Average Trading Frequency	Lagged Trade Frequency	Current Trade Volume	Prior Day Rolling Average Trading Volume	Lagged Trading Volume	Adjusted R ²	Residual Standard Error	
HAR_M		*	***	*	***	***	***	***	0.4005	0.4440	0.7451
HAR_norol			***	***	***	***	***	***	0.4246	0.4391	0.7488
HAR_TP			***	*	***	***	***	***	0.3305	0.3271	0.6203
HAR_TV			***	***	***	***	*	***	0.3090	0.6363	0.8329
HAR_TVTF			***	***	***	***	***	***	0.4456	0.4413	0.7474
HAR_TVTF_norol			***	***	***	***	***	***	0.4418	0.439	0.749

Symbol α Level

*** 0.001

** 0.01

* 0.05

0.1

DIS RV Model	Lagged Volatility	Prior Day Rolling Average Volatility	Current Trading Frequency	Prior Day Rolling Average Trading Frequency	Lagged Trade Frequency	Current Trade Volume	Prior Day Rolling Average Trading Volume	Lagged Trading Volume	Adjusted R ²	Residual Standard Error	
HAR_M		**	***	**	***	***	***	***	0.4419	0.4382	0.6981
HAR_norol			***	***	***	***	***	***	0.4295	0.4258	0.7115
HAR_TP			***	***	***	***	***	***	0.3136	0.3161	0.7055
HAR_TV			***	***	***	***	***	***	0.2965	0.2929	0.6282
HAR_TVTF			***	***	***	***	***	***	0.4315	0.4272	0.7031
HAR_TVTF_norol			***	***	***	***	***	***	0.4283	0.4264	0.7161

Symbol α Level

*** 0.001

** 0.01

* 0.05

0.1



Ford Motor Company (F)

F Range4 Model	Lagged Volatility	Prior Day Rolling Average Volatility	Current Trading Frequency	Prior Day Rolling Average Trading Frequency	Lagged Trade Frequency	Current Trade Volume	Prior Day Rolling Average Trading Volume	Lagged Trading Volume	R ²	Adjusted R ²	Residual Standard Error
HAR_M		**	***	*	***	***	***	*	0.6620	0.6594	
HAR_norol	**		***	***	***	***	***	*	0.6555	0.6533	
HAR_TP	**		***	***	***	***	***	*	0.6205	0.6296	
HAR_TV	***		***	***	***	***	***	*	0.4820	0.4903	
HAR_TVTF			***	***	***	***	***	*	0.6534	0.6508	
HAR_TVTF_norol			***	***	***	***	***	**	0.5133	0.5109	

F Range5 Model	Lagged Volatility	Prior Day Rolling Average Volatility	Current Trading Frequency	Prior Day Rolling Average Trading Frequency	Lagged Trade Frequency	Current Trade Volume	Prior Day Rolling Average Trading Volume	Lagged Trading Volume	R ²	Adjusted R ²	Residual Standard Error
HAR_M			***	***	***	***	***	**	0.5226	0.516	
HAR_norol			***	***	***	***	***	**	0.5163	0.5132	
HAR_TP			***	***	***	***	***	**	0.4736	0.471	
HAR_TV			***	***	*	***	***	***	0.3891	0.396	
HAR_TVTF			***	***	**	***	***	**	0.5169	0.5132	
HAR_TVTF_norol			***	***	**	***	***	**	0.5133	0.5109	

F RV Model	Lagged Volatility	Prior Day Rolling Average Volatility	Current Trading Frequency	Prior Day Rolling Average Trading Frequency	Lagged Trade Frequency	Current Trade Volume	Prior Day Rolling Average Trading Volume	Lagged Trading Volume	Adjusted R ²	Residual Standard Error
HAR_M			***	***	*	***	***	**	0.5536	0.5793
HAR_norol			***	***	**	***	***	**	0.5792	0.5765
HAR_TP			***	***	*	***	***	**	0.5162	0.5137
HAR_TV			***	***	**	***	***	*	0.433	0.4302
HAR_TVTF			***	***	**	***	***	*	0.5009	0.5777
HAR_TVTF_norol			***	***	**	***	***	**	0.5769	0.5746

Symbol

α Level

0.00

0.01

0.05

0.10



Home Depot (HD)

HD Range4 Model	Lagged Volatility	Prior Day Rolling Average Volatility	Current Trading Frequency	Prior Day Rolling Average Trading Frequency	Lagged Trade Frequency	Current Trade Volume	Prior Day Rolling Average Trading Volume	Lagged Trading Volume	Adjusted R ²	Residual Standard Error	
HAR_M			***	***	**	***	***	***	0.5438	0.5389	0.679
HAR_norol	*		***	***	***	***	***	***	0.5347	0.5317	0.6843
HAR_TP			***	***	***	***	***	*	0.4501	0.4272	0.7560
HAR_TV			***	***	***	***	***	*	0.3625	0.3793	0.7879
HAR_TVTF			***	***	**	***	***	***	0.5409	0.5374	0.6802
HAR_TVTF_norol			***	***	**	***	***	***	0.531	0.5286	0.6868

Symbol

▲ Level

HD Range4 Model	Lagged Volatility	Prior Day Rolling Average Volatility	Current Trading Frequency	Prior Day Rolling Average Trading Frequency	Lagged Trade Frequency	Current Trade Volume	Prior Day Rolling Average Trading Volume	Lagged Trading Volume	Adjusted R ²	Residual Standard Error	
HAR_M			***	**	*	***	***	***	0.4997	0.4542	0.7309
HAR_norol			***	***	**	***	***	***	0.4537	0.4503	0.7414
HAR_TP	**		***	***	***	***	***	***	0.3575	0.3542	0.8036
HAR_TV			***	***	***	***	***	***	0.3371	0.3327	0.8163
HAR_TVTF			***	*	*	***	***	***	0.6023	0.6041	0.7388
HAR_TVTF_norol			***	***	**	***	***	***	0.4536	0.4500	0.7411

Symbol

▲ Level

HD RV Model	Lagged Volatility	Prior Day Rolling Average Volatility	Current Trading Frequency	Prior Day Rolling Average Trading Frequency	Lagged Trade Frequency	Current Trade Volume	Prior Day Rolling Average Trading Volume	Lagged Trading Volume	Adjusted R ²	Residual Standard Error	
HAR_M			***	**	**	***	***	***	0.4247	0.4189	0.7823
HAR_norol			***	***	***	***	***	***	0.4131	0.4094	0.7885
HAR_TP			***	***	***	***	***	***	0.3141	0.3106	0.8303
HAR_TV	*		***	*	*	***	*	***	0.2644	0.2607	0.8996
HAR_TVTF			***	*	**	***	***	***	0.4238	0.4194	0.782
HAR_TVTF_norol			***	***	***	***	***	***	0.413	0.4101	0.7893

Symbol

▲ Level



Lowe's (LOW)

LOW Range/Model	Lagged Volatility	Prior Day Rolling Average Volatility	Current Trading Frequency	Prior Day Rolling Average Trading Frequency	Lagged Trade Frequency	Current Trade Volume	Prior Day Rolling Average Trading Volume	Lagged Trading Volume	R ²	Adjusted R ²	Residual Standard Error
HAR_M		***	***	***	***	***	***	***	0.5379	0.5332	0.6832
HAR_Mvol	***		***	***	***	***	***	***	0.5252	0.5222	0.6915
HAR_TP			***	***	***	***	***	***	0.4171	0.4144	0.7055
HAR_TV	**					***	**	***	0.3247	0.3212	0.8239
HAR_TVTF			***	*		***		***	0.5176	0.5139	0.6972
HAR_TVTF_Mvol			***		*	***		***	0.5113	0.5105	0.6996

Symbol

α Level

***0.001

**0.01

*0.05

.0.1

LOW Range/Model	Lagged Volatility	Prior Day Rolling Average Volatility	Current Trading Frequency	Prior Day Rolling Average Trading Frequency	Lagged Trade Frequency	Current Trade Volume	Prior Day Rolling Average Trading Volume	Lagged Trading Volume	Adjusted R ²	Residual Standard Error	
HAR_M		**			***	***	***	***	0.4043	0.3982	0.7757
HAR_Mvol			***		*	***	***	***	0.3944	0.3906	0.7807
HAR_TP			***		*	***	***	***	0.2633	0.2787	0.8487
HAR_TV						***	*	***	0.229	0.2251	0.8803
HAR_TVTF			***		***	***	***	***	0.3942	0.3895	0.7815
HAR_TVTF_Mvol			***		**	***	***	***	0.3922	0.3891	0.7816

Symbol

α Level

***0.001

**0.01

*0.05

.0.1

LOW RV Model	Lagged Volatility	Prior Day Rolling Average Volatility	Current Trading Frequency	Prior Day Rolling Average Trading Frequency	Lagged Trade Frequency	Current Trade Volume	Prior Day Rolling Average Trading Volume	Lagged Trading Volume	Adjusted R ²	Residual Standard Error	
HAR_M		***	***	***	***	***	***	***	0.5011	0.499	0.71
HAR_Mvol	***		***	***	***	***	***	***	0.4324	0.4301	0.721
HAR_TP	*		***	***	***	***	***	***	0.3652	0.3619	0.7888
HAR_TV						*	***	***	0.2253	0.2213	0.8524
HAR_TVTF			***	***	***	***	***	***	0.4945	0.4906	0.7137
HAR_TVTF_Mvol			***	***	***	***	***	***	0.4912	0.4765	0.7221

Symbol

α Level

***0.001

**0.01

*0.05

.0.1

McDonald's Corp. (MCD)

MCD Model	Lagged Volatility	Prior Day Rolling Average Volatility	Current Trading Frequency	Prior Day Rolling Average Trading Frequency	Lagged Trade Frequency	Current Trade Volume	Prior Day Rolling Average Trading Volume	Lagged Trading Volume	Adjusted R ²	Residual Standard Error	
HAR_M		***	***	***	***	***	***	***	0.5099	0.5088	0.6428
HAR_Mvol	***		***	***	***	***	***	***	0.5716	0.5689	0.6565
HAR_TP			***	***	***	***	***	***	0.4889	0.4843	0.7181
HAR_TV						***	***	***	0.4582	0.4555	0.7379
HAR_TVTF			***	***	***	***	***	***	0.5751	0.5718	0.6544
HAR_TVTF_Mvol			***	***	***	***	***	***	0.5581	0.5555	0.6564

Symbol

**

*

.

α Level

0.001

0.01

0.05

0.1

MCD Range/Model	Lagged Volatility	Prior Day Rolling Average Volatility	Current Trading Frequency	Prior Day Rolling Average Trading Frequency	Lagged Trade Frequency	Current Trade Volume	Prior Day Rolling Average Trading Volume	Lagged Trading Volume	Adjusted R ²	Residual Standard Error	
HAR_M		***	***	***	***	***	***	***	0.4613	0.4558	0.7377
HAR_Mvol	*		***	***	***	***	***	***	0.4448	0.4412	0.7475
HAR_TP			***	***	***	***	***	***	0.3877	0.3845	0.7972
HAR_TV					***	***	***	***	0.3721	0.3689	0.7844
HAR_TVTF			***	***	***	***	***	***	0.4512	0.447	0.7436
HAR_TVTF_Mvol			***	***	***	***	***	***	0.4413	0.4385	0.7486

Symbol

**

*

.

α Level

0.001

0.01

0.05

0.1

MCD RV Model	Lagged Volatility	Prior Day Rolling Average Volatility	Current Trading Frequency	Prior Day Rolling Average Trading Frequency	Lagged Trade Frequency	Current Trade Volume	Prior Day Rolling Average Trading Volume	Lagged Trading Volume	Adjusted R ²	Residual Standard Error	
HAR_M		***	***	***	***	***	***	***	0.4629	0.4574	0.7388
HAR_Mvol		*	***	***	***	***	***	***	0.4403	0.4368	0.7505
HAR_TP	***		***	***	***	***	***	***	0.3799	0.3768	0.7894
HAR_TV	***				***	***	***	***	0.3611	0.3578	0.8014
HAR_TVTF			***	***	***	***	***	***	0.4577	0.4538	0.7362
HAR_TVTF_Mvol			***	***	***	***	***	***	0.4483	0.4375	0.73

Symbol

**

*

.

α Level

0.001

0.01

0.05

0.1

Nike, Inc. (NKE)

NKE Range/ Model	Lagged Volatility	Prior Day Rolling Average Volatility	Current Trading Frequency	Prior Day Rolling Average Trading Frequency	Lagged Trade Frequency	Current Trade Volume	Prior Day Rolling Average Trading Volume	Lagged Trading Volume	Adjusted R ²	Residual Standard Error	
HAR_M	*	***	***	*	***	***	***	***	0.4839	0.4787	0.722
HAR_Mvol	***		***	***	***	***	***	***	0.4409	0.4374	0.7501
HAR_TP	*		***	***	***	***	***	***	0.3608	0.3574	0.8017
HAR_TV	***					***	***	***	0.4112	0.4083	0.7688
HAR_TVTF			***	***	*	***	*	***	0.4374	0.4331	0.7528
HAR_TVTF_Mvol			***	***	***	***	***	***	0.3968	0.3937	0.7788

Symbol

α Level

**

*

.

0.001

0.01

0.05

0.1

NKE Rangess Model	Lagged Volatility	Prior Day Rolling Average Volatility	Current Trading Frequency	Prior Day Rolling Average Trading Frequency	Lagged Trade Frequency	Current Trade Volume	Prior Day Rolling Average Trading Volume	Lagged Trading Volume	Adjusted R ²	Residual Standard Error	
HAR_M		***	***	***	**	**	***	***	0.41	0.404	0.772
HAR_Mvol	**		**	***	***	***	***	***	0.3801	0.3761	0.7899
HAR_TP			***	***	***	**	***	***	0.2919	0.2883	0.8426
HAR_TV						***	***	***	0.362	0.3568	0.8088
HAR_TVTF			***	**	**	**	***	***	0.3894	0.3848	0.7844
HAR_TVTF_Mvol			***	***	*	***	***	***	0.3871	0.3836	0.7875

Symbol

α Level

**

*

.

0.001

0.01

0.05

0.1

NKE RV Model	Lagged Volatility	Prior Day Rolling Average Volatility	Current Trading Frequency	Prior Day Rolling Average Trading Frequency	Lagged Trade Frequency	Current Trade Volume	Prior Day Rolling Average Trading Volume	Lagged Trading Volume	Adjusted R ²	Residual Standard Error	
HAR_M		***	**		***	*	*	***	0.38	0.3737	0.7814
HAR_Mvol		*		***	***	*	***	***	0.3485	0.3453	0.8091
HAR_TP			**	***	***		***	***	0.2875	0.2838	0.856
HAR_TV						*	***	***	0.3465	0.3422	0.8111
HAR_TVTF			**	*	**	*	***	***	0.3616	0.3567	0.8021
HAR_TVTF_Mvol			*		***	*	***	***	0.34	0.3367	0.8143

Symbol

α Level

**

*

.

0.001

0.01

0.05

0.1

Starbucks Corp. (SBUX)

SBUX All Model	Lagged Volatility	Prior Day Rolling Average Volatility	Current Trading Frequency	Prior Day Rolling Average Trading Frequency	Lagged Trade Frequency	Current Trade Volume	Prior Day Rolling Average Trading Volume	Lagged Trading Volume	R ²	Adjusted R ²	Residual Standard Error
HAR_M		***		***	**	***			0.4684	0.464	0.7321
HAR_Mvol									0.4542	0.4505	0.7411
HAR_TP			**						0.426	0.4352	0.7515
HAR_TV	*					***		***	0.4386	0.437	0.7501
HAR_TVTF				*	***				0.4576	0.4534	0.7393
HAR_TVTF_norol					**	***			0.4569	0.4481	0.7429

Symbol

α Level

0.001

**

0.01

*

0.05

0.1

SBUX Range's Model	Lagged Volatility	Prior Day Rolling Average Volatility	Current Trading Frequency	Prior Day Rolling Average Trading Frequency	Lagged Trade Frequency	Current Trade Volume	Prior Day Rolling Average Trading Volume	Lagged Trading Volume	R ²	Adjusted R ²	Residual Standard Error
HAR_M		*		**		***		*	0.4001	0.394	0.7755
HAR_Mvol					**	***			0.3955	0.3946	0.7545
HAR_TP			**						0.379	0.3759	0.79
HAR_TV						***		***	0.3711	0.3679	0.795
HAR_TVTF				*	***		*		0.3952	0.3906	0.7884
HAR_TVTF_norol					**	***			0.3883	0.3852	0.7841

Symbol

α Level

0.001

**

0.01

*

0.05

0.1

SBUX RV Model	Lagged Volatility	Prior Day Rolling Average Volatility	Current Trading Frequency	Prior Day Rolling Average Trading Frequency	Lagged Trade Frequency	Current Trade Volume	Prior Day Rolling Average Trading Volume	Lagged Trading Volume	R ²	Adjusted R ²	Residual Standard Error
HAR_M						*			0.273	0.2656	0.857
HAR_Mvol									0.2584	0.2547	0.8833
HAR_TP			**						0.27	0.2683	0.8565
HAR_TV									0.2904	0.2485	0.866
HAR_TVTF									0.271	0.2655	0.857
HAR_TVTF_norol									0.2593	0.2556	0.8623

Symbol

α Level

0.001

**

0.01

*

0.05

0.1

Target Corporation (TGT)

TGT Range's Model	Lagged Volatility	Prior Day Rolling Average Volatility	Current Trading Frequency	Prior Day Rolling Average Trading Frequency	Lagged Trade Frequency	Current Trade Volume	Prior Day Rolling Average Trading Volume	Lagged Trading Volume	R ²	Adjusted R ²	Residual Standard Error
HAR_M		***	***	***	***	***	*	***	0.5261	0.5212	0.6919
HAR_Mvol	***		***	***	***	***		***	0.5068	0.5025	0.7053
HAR_TP			***	***					0.4164	0.4274	0.7668
HAR_TV	**					***	*	***	0.3735	0.3706	0.7834
HAR_TVTF			***	***	**	***	**	***	0.5076	0.5039	0.7044
HAR_TVTF_norol			***	***	***	***	***	***	0.4845	0.4819	0.7189

Symbol

α Level

0.001

**

0.01

*

0.05

0.1

TGT Range's Model	Lagged Volatility	Prior Day Rolling Average Volatility	Current Trading Frequency	Prior Day Rolling Average Trading Frequency	Lagged Trade Frequency	Current Trade Volume	Prior Day Rolling Average Trading Volume	Lagged Trading Volume	R ²	Adjusted R ²	Residual Standard Error
HAR_M		***	***	***	***	***	*	***	0.452	0.4464	0.744
HAR_Mvol			***	***	***	***		***	0.4272	0.4235	0.7589
HAR_TP			***	***	*				0.3371	0.3343	0.8166
HAR_TV						***	*	***	0.3012	0.2976	0.8381
HAR_TVTF			***	***	***	***	*	***	0.4448	0.4405	0.748
HAR_TVTF_norol			***	***	***	***	***	***	0.4256	0.4237	0.7589

Symbol

α Level

0.001

**

0.01

*

0.05

0.1

TGT RV Model	Lagged Volatility	Prior Day Rolling Average Volatility	Current Trading Frequency	Prior Day Rolling Average Trading Frequency	Lagged Trade Frequency	Current Trade Volume	Prior Day Rolling Average Trading Volume	Lagged Trading Volume	R ²	Adjusted R ²	Residual Standard Error
HAR_M		***	***	***	***	***		***	0.5176	0.5127	0.6981
HAR_Mvol			***	***	***	***		***	0.4959	0.4937	0.7165
HAR_TP			***	***					0.3839	0.3806	0.7865
HAR_TV						**	**	***	0.3175	0.314	0.8282
HAR_TVTF			***	**	**	***	*	***	0.5094	0.5057	0.7031
HAR_TVTF_norol			***	***	***	***	***	***	0.4888	0.4872	0.7181

Symbol

α Level

0.001

**

0.01

*

0.05

0.1

Time Warner Inc. (TWX)

TWIX RV Model	Lagged Volatility	Prior Day Rolling Average Volatility	Current Trading Frequency	Prior Day Rolling Average Trading Frequency	Lagged Trade Frequency	Current Trade Volume	Prior Day Rolling Average Trading Volume	Lagged Trading Volume	Adjusted R ²	Residual Standard Error	
HAR_M	***	***	***	***	***	***	***	***	0.4432	0.4375	0.75
HAR_Mvol	***	***	***	***	**	***	***	***	0.4333	0.4297	0.7562
HAR_TP	***	***	***	***	*				0.3267	0.3233	0.8226
HAR_TV	***	***	***	***	*	*		***	0.3091	0.3056	0.8333
HAR_TVTF		***	***	*	*	***	*	***	0.4097	0.4052	0.7712
HAR_TVTF_norol		***	***		*	***		***	0.4009	0.3979	0.7715

Symbol

α Level

0.001

**

0.01

*

0.05

0.1

TWIX Range's Model	Lagged Volatility	Prior Day Rolling Average Volatility	Current Trading Frequency	Prior Day Rolling Average Trading Frequency	Lagged Trade Frequency	Current Trade Volume	Prior Day Rolling Average Trading Volume	Lagged Trading Volume	Adjusted R ²	Residual Standard Error	
HAR_M	***	*	***	***	***	***		***	0.371	0.3656	0.7985
HAR_Mvol	***		***	***	**	***		***	0.366	0.362	0.7987
HAR_TP	***		*						0.2855	0.2816	0.8592
HAR_TV	***		***		**	***		***	0.2768	0.2731	0.8526
HAR_TVTF			***			***		***	0.341	0.335	0.8146
HAR_TVTF_norol			***		**	***		***	0.3384	0.335	0.8155

Symbol

α Level

0.001

**

0.01

*

0.05

0.1

TWIX RV Model	Lagged Volatility	Prior Day Rolling Average Volatility	Current Trading Frequency	Prior Day Rolling Average Trading Frequency	Lagged Trade Frequency	Current Trade Volume	Prior Day Rolling Average Trading Volume	Lagged Trading Volume	Adjusted R ²	Residual Standard Error	
HAR_M	***	***	*	*	*	*		***	0.3274	0.3206	0.8243
HAR_Mvol	***	***	*	*	*	**		***	0.3216	0.3173	0.8293
HAR_TP	***	***		*	*	*		***	0.1953	0.1912	0.8899
HAR_TV	***	***			*	*	*	***	0.1807	0.1765	0.9075
HAR_TVTF		***	*		*	*	*	***	0.2897	0.2843	0.846
HAR_TVTF_norol		***	*		*	**		***	0.286	0.2823	0.8471

Symbol

α Level

0.001

**

0.01

*

0.05

0.1

Appendix C

Full VAR Results

Amazon.com (AMZN)

Model	Response Predictors	Current Volatility	Current Trading Volume	Prior Day Rolling Average Trading Volume	Current Volatility	Current Trading Volume	Prior Day Rolling Average Trading Volume	R ²	Adjusted R ²	Residual Standard Error
VAR_TV	Current Volatility (RA)	***			0.1378	0.2039	0.9518	0.1773	0.1742	0.9068
	Current Trading Volume		***		-0.1186	0.4814	0.1023	0.256	0.2631	0.8631
	Prior Day Rolling Average Trading Volume			***	-0.0337	0.1459	0.9486	0.9759	0.9758	0.1556
VAR_TF	Current Volatility (RA)	***			0.1241	0.3309	0.9092	0.1849	0.1818	0.9026
	Current Trading Frequency		***		-0.0734	0.5436	0.0888	0.2767	0.274	0.8009
	Prior Day Rolling Average Trading Frequency			***	-0.0272	0.1363	0.9499	0.9749	0.9748	0.1588
VAR_TV	Current Volatility (RS)	***			0.0849	0.2731	0.9396	0.122	0.1187	0.9358
	Current Trading Volume		***		-0.0057	0.4717	0.1023	0.2568	0.263	0.8632
	Prior Day Rolling Average Trading Volume			***	-0.0283	0.1411	0.9479	0.9758	0.9757	0.156
VAR_TF	Current Volatility (RS)	***			0.0695	0.3112	-0.0024	0.1305	0.1272	0.9311
	Current Trading Frequency		***		-0.0593	0.5314	0.0890	0.276	0.2732	0.8013
	Prior Day Rolling Average Trading Frequency			***	-0.0237	0.1330	0.9495	0.9748	0.9747	0.159
VAR_TV	Current Volatility (RV)	***			-0.0585	0.3243	-0.0025	0.09007	0.07657	0.9601
	Current Trading Volume		***		-0.1488	0.5863	0.0963	0.2655	0.2627	0.8575
	Prior Day Rolling Average Trading Volume			***	0.0059	0.1172	0.9498	0.9754	0.9753	0.1574
VAR_TF	Current Volatility (RV)	***			-0.0591	0.3421	-0.0024	0.08713	0.08366	0.9584
	Current Trading Frequency		***		-0.1810	0.6268	0.0714	0.2895	0.2868	0.8433
	Prior Day Rolling Average Trading Frequency			***	0.0128	0.1074	0.9527	0.9746	0.9745	0.1588

Symbol is Level

*** 0.001

** 0.01

* 0.05

0.1

The Walt Disney Company (DIS)

Model	Response Predictors	Current Volatility	Current Trading Volume	Prior Day Rolling Average Trading Volume	Current Volatility	Current Trading Volume	Prior Day Rolling Average Trading Volume	R ²	Adjusted R ²	Residual Standard Error
VAR_TV	Current Volatility (RA)	***			0.0287	0.3287	-0.0481	0.1773	0.1742	0.9068
	Current Trading Volume		***		0.0775	0.4187	0.1329	0.2560	0.2531	0.8631
	Prior Day Rolling Average Trading Volume			***	0.0147	0.0823	0.9602	0.9759	0.9758	0.1556
VAR_TF	Current Volatility (RA)	***			-0.0529	0.5180	-0.0496	0.1849	0.1818	0.9026
	Current Trading Frequency		***		-0.0241	0.6092	0.0789	0.2767	0.2740	0.8009
	Prior Day Rolling Average Trading Frequency			***	0.0067	0.0928	0.9597	0.9749	0.9748	0.1588
VAR_TV	Current Volatility (RS)	***			0.0099	0.3438	-0.0423	0.1220	0.1187	0.9358
	Current Trading Volume		***		0.0222	0.4169	0.1327	0.2558	0.2530	0.8632
	Prior Day Rolling Average Trading Volume			***	0.0096	0.0839	0.9604	0.9758	0.9757	0.1580
VAR_TF	Current Volatility (RS)	***			-0.0775	0.5409	-0.0487	0.1305	0.1272	0.9311
	Current Trading Frequency		***		-0.0949	0.6112	0.0786	0.2760	0.2732	0.8013
	Prior Day Rolling Average Trading Frequency			***	-0.0037	0.0862	0.9594	0.9746	0.9747	0.1592
VAR_TV	Current Volatility (RV)	***			0.1047	0.3446	-0.0498	0.0801	0.0766	0.9601
	Current Trading Volume		***		0.0387	0.4098	0.1319	0.2655	0.2627	0.8575
	Prior Day Rolling Average Trading Volume			***	0.0130	0.0823	0.9602	0.9754	0.9753	0.1574
VAR_TF	Current Volatility (RV)	***			-0.0028	0.3536	-0.0511	0.0871	0.0837	0.9584
	Current Trading Frequency		***		-0.0223	0.6063	0.0787	0.2895	0.2868	0.8433
	Prior Day Rolling Average Trading Frequency			***	-0.0019	-0.0975	0.9584	0.9746	0.9745	0.1588

Symbol is Level

*** 0.001

** 0.01

* 0.05

0.1

Ford Motor Company (F)

Model	Response Predictors	Current Volatility	Current Trading Volume	Prior Day Rolling Average Trading Volume	Current Volatility	Current Trading Volume	Prior Day Rolling Average Trading Volume	R ²	Adjusted R ²	Residual Standard Error
VAR_TV	Current Volatility (RA)	***	***	***	0.1107	0.4710	0.0214	0.3152	0.3126	0.6288
	Current Trading Volume	***	***	***	-0.0352	0.6886	0.1242	0.4194	0.4172	0.7635
	Prior Day Rolling Average Trading Volume	***	***	***	-0.0084	0.1138	0.9409	0.9873	0.9873	0.1129
VAR_TF	Current Volatility (RA)	***	***	***	0.0076	0.4074	0.0075	0.282	0.2793	0.8467
	Current Trading Volume	***	***	***	-0.0832	0.6158	0.1441	0.3938	0.3915	0.7792
	Prior Day Rolling Average Trading Volume	***	***	***	-0.0137	0.1149	0.9444	0.9873	0.9873	0.1129
VAR_TV	Current Volatility (RS)	***	***	***	0.0888	0.4759	-0.0034	0.2888	0.2841	0.848
	Current Trading Volume	***	***	***	0.0135	0.6773	0.1236	0.4188	0.4146	0.7639
	Prior Day Rolling Average Trading Volume	***	***	***	-0.0052	0.1171	0.9409	0.9873	0.9872	0.113
VAR_TF	Current Volatility (RS)	***	***	***	0.0777	0.4084	0.0425	0.2347	0.2318	0.8793
	Current Trading Volume	***	***	***	0.0218	0.6324	0.1470	0.3914	0.3891	0.7808
	Prior Day Rolling Average Trading Volume	***	***	***	-0.0103	0.1110	0.9447	0.9873	0.9873	0.1129
VAR_TV	Current Volatility (RV)	***	***	***	0.0645	0.5395	-0.0021	0.3348	0.332	0.8156
	Current Trading Volume	***	***	***	0.0055	0.6664	0.1228	0.4182	0.417	0.7637
	Prior Day Rolling Average Trading Volume	***	***	***	0.0014	0.1133	0.9408	0.9873	0.9872	0.1131
VAR_TF	Current Volatility (RV)	***	***	***	0.0559	0.4473	-0.0116	0.2644	0.2616	0.8578
	Current Trading Volume	***	***	***	0.0241	0.6303	0.1487	0.3915	0.3891	0.7807
	Prior Day Rolling Average Trading Volume	***	***	***	-0.0062	0.1083	0.9449	0.9873	0.9872	0.1131



Home Depot (HD)

Model	Response Predictors	Current Volatility	Current Trading Volume	Prior Day Rolling Average Trading Volume	Current Volatility	Current Trading Volume	Prior Day Rolling Average Trading Volume	R ²	Adjusted R ²	Residual Standard Error
VAR_TV	Current Volatility (RA)	***	***	***	0.0170	0.5681	0.0141	0.3395	0.333	0.8151
	Current Trading Volume	***	***	***	-0.0502	0.3853	0.1114	0.1616	0.1584	0.9166
	Prior Day Rolling Average Trading Volume	***	***	***	0.0071	0.0940	0.9599	0.9607	0.9606	0.1805
VAR_TF	Current Volatility (RA)	***	***	***	-0.1189	0.6279	-0.0375	0.3096	0.306	0.8215
	Current Trading Volume	***	***	***	-0.1645	0.6566	0.0961	0.3742	0.3718	0.7907
	Prior Day Rolling Average Trading Volume	***	***	***	-0.0016	0.0859	0.9623	0.9785	0.9785	0.1487
VAR_TV	Current Volatility (RS)	***	***	***	-0.0486	0.5656	0.0146	0.3047	0.3021	0.8356
	Current Trading Volume	***	***	***	0.0033	0.3750	0.1107	0.1605	0.1573	0.9175
	Prior Day Rolling Average Trading Volume	***	***	***	0.0028	0.0959	0.9599	0.9607	0.9605	0.1805
VAR_TF	Current Volatility (RS)	***	***	***	-0.1023	0.5925	-0.0348	0.2673	0.2645	0.8578
	Current Trading Volume	***	***	***	-0.1335	0.6319	0.0976	0.3695	0.3671	0.7936
	Prior Day Rolling Average Trading Volume	***	***	***	-0.0075	0.0870	0.9621	0.9786	0.9785	0.1488
VAR_TV	Current Volatility (RV)	***	***	***	-0.0899	0.5079	-0.0140	0.2284	0.2254	0.8877
	Current Trading Volume	***	***	***	-0.0182	0.3666	0.1091	0.1598	0.1566	0.9176
	Prior Day Rolling Average Trading Volume	***	***	***	0.0100	0.0934	0.9598	0.9607	0.9606	0.1804
VAR_TF	Current Volatility (RV)	***	***	***	-0.1868	0.5382	-0.0668	0.2081	0.2031	0.8931
	Current Trading Volume	***	***	***	-0.1279	0.6275	0.0929	0.3688	0.3664	0.7941
	Prior Day Rolling Average Trading Volume	***	***	***	0.0024	0.0816	0.9625	0.9786	0.9785	0.1487



Lowe's (LOW)

Model	Response Predictors	Current Volatility	Current Trading Volume	Prior Day Rolling Average Trading Volume	Current Volatility	Current Trading Volume	Prior Day Rolling Average Trading Volume	R ²	Adjusted R ²	Residual Standard Error
VAR_TV	Current Volatility (RA)	***	***	***	0.1447	0.4454	-0.0054	0.3894	0.2666	0.8534
	Current Trading Volume	***	***	***	-0.016	0.3052	0.2483	0.2095	0.2038	0.893
	Prior Day Rolling Average Trading Volume	***	***	***	0.0106	0.0851	0.959	0.9607	0.9628	0.1319
VAR_TF	Current Volatility (RA)	***	***	***	0.0419	0.446	0.0059	0.2077	0.2047	0.8774
	Current Trading Volume	***	***	***	-0.0733	0.5317	0.1741	0.3353	0.3328	0.8172
	Prior Day Rolling Average Trading Volume	***	***	***	-0.0097	0.1032	0.9497	0.9842	0.9842	0.1259
VAR_TV	Current Volatility (RS)	***	***	***	0.0309	0.4321	-0.0115	0.1934	0.1903	0.8981
	Current Trading Volume	***	***	***	-0.0036	0.3014	0.2476	0.2064	0.2034	0.8931
	Prior Day Rolling Average Trading Volume	***	***	***	0.0064	0.104	0.9593	0.9606	0.9609	0.1321
VAR_TF	Current Volatility (RS)	***	***	***	-0.036	0.3895	0.0148	0.1439	0.1408	0.9202
	Current Trading Volume	***	***	***	-0.0555	0.5123	0.1754	0.3338	0.3313	0.8181
	Prior Day Rolling Average Trading Volume	***	***	***	-0.0129	0.1039	0.9497	0.9843	0.9842	0.1257
VAR_TV	Current Volatility (RV)	***	***	***	-0.0465	0.4251	-0.0684	0.1526	0.1493	0.9229
	Current Trading Volume	***	***	***	0.0002	0.2985	0.2479	0.2064	0.2034	0.8931
	Prior Day Rolling Average Trading Volume	***	***	***	-0.011	0.0817	0.9572	0.9607	0.9607	0.1319
VAR_TF	Current Volatility (RV)	***	***	***	-0.1447	0.4356	-0.0742	0.1198	0.1165	0.9404
	Current Trading Volume	***	***	***	-0.121	0.5608	0.158	0.3409	0.3384	0.8138
	Prior Day Rolling Average Trading Volume	***	***	***	-0.0033	0.0991	0.9497	0.9842	0.9841	0.1261



McDonald's Corp. (MCD)

Model	Response Predictors	Current Volatility	Current Trading Volume	Prior Day Rolling Average Trading Volume	Current Volatility	Current Trading Volume	Prior Day Rolling Average Trading Volume	R ²	Adjusted R ²	Residual Standard Error
VAR_TV	Current Volatility (RA)	***	***	***	0.0136	0.5889	0.0331	0.4122	0.4099	0.7205
	Current Trading Volume	***	***	***	-0.0649	0.4247	0.1229	0.1991	0.196	0.8813
	Prior Day Rolling Average Trading Volume	***	***	***	0.02113	0.0783	0.9627	0.9721	0.972	0.1673
VAR_TF	Current Volatility (RA)	***	***	***	-0.0823	0.6288	-0.0548	0.3688	0.3641	0.7108
	Current Trading Volume	***	***	***	-0.2044	0.7122	0.1086	0.4349	0.4327	0.7367
	Prior Day Rolling Average Trading Volume	***	***	***	0.001	0.0792	0.962	0.9835	0.9834	0.1287
VAR_TV	Current Volatility (RS)	***	***	***	-0.031	0.5564	0.043	0.3553	0.3528	0.7484
	Current Trading Volume	***	***	***	-0.0801	0.4178	0.1229	0.1989	0.1958	0.8814
	Prior Day Rolling Average Trading Volume	***	***	***	0.0104	0.0844	0.9635	0.9719	0.9718	0.168
VAR_TF	Current Volatility (RS)	***	***	***	-0.0865	0.5587	0.0448	0.29	0.2873	0.7953
	Current Trading Volume	***	***	***	-0.1857	0.6839	0.1148	0.4277	0.4255	0.7413
	Prior Day Rolling Average Trading Volume	***	***	***	-0.0068	0.0841	0.9615	0.9835	0.9835	0.1286
VAR_TV	Current Volatility (RV)	***	***	***	-0.1102	0.5815	-0.0498	0.3187	0.3161	0.7815
	Current Trading Volume	***	***	***	-0.0273	0.4682	0.1185	0.1965	0.1935	0.8657
	Prior Day Rolling Average Trading Volume	***	***	***	0.0163	0.0818	0.9643	0.972	0.9719	0.1678
VAR_TF	Current Volatility (RV)	***	***	***	-0.1975	0.6146	-0.1303	0.2748	0.2721	0.8062
	Current Trading Volume	***	***	***	-0.1479	0.6823	0.1047	0.4255	0.4234	0.7428
	Prior Day Rolling Average Trading Volume	***	***	***	-0.0006	0.0806	0.9618	0.9835	0.9834	0.1287



Nike, Inc. (NKE)

Model	Response Predictors	Current Volatility	Current Trading Volume	Prior Day Rolling Average Trading Volume	Current Volatility	Current Trading Volume	Prior Day Rolling Average Trading Volume	R ²	Adjusted R ²	Residual Standard Error
VAR_TV	Current Volatility (RA)	***	***	***	0.1423	0.6006	-0.1472	0.3885	0.3841	0.7833
	Current Trading Volume	***	***	***	-0.0534	0.4249	0.2272	0.295	0.2903	0.8416
	Prior Day Rolling Average Trading Volume	***	***	***	0.0007	0.0867	0.9648	0.9871	0.9871	0.1137
VAR_TF	Current Volatility (RA)	***	***	***	0.1037	0.607	-0.2897	0.3035	0.3008	0.8348
	Current Trading Volume	***	***	***	-0.0791	0.6197	0.2197	0.5548	0.5531	0.6688
	Prior Day Rolling Average Trading Volume	***	***	***	-0.0093	0.0714	0.954	0.9938	0.9937	0.0792
VAR_TV	Current Volatility (RS)	***	***	***	0.0507	0.6181	-0.183	0.34542	0.3427	0.8105
	Current Trading Volume	***	***	***	-0.0635	0.4264	0.2289	0.2961	0.2935	0.8409
	Prior Day Rolling Average Trading Volume	***	***	***	-0.0018	0.0875	0.9648	0.9871	0.9871	0.1137
VAR_TF	Current Volatility (RS)	***	***	***	0.0241	0.6084	-0.2818	0.2943	0.2941	0.8601
	Current Trading Volume	***	***	***	-0.0831	0.6163	0.2213	0.5556	0.5539	0.6682
	Prior Day Rolling Average Trading Volume	***	***	***	-0.0101	0.0711	0.9542	0.9938	0.9937	0.0791
VAR_TV	Current Volatility (RV)	***	***	***	0.0487	0.6115	-0.1777	0.3321	0.3295	0.8171
	Current Trading Volume	***	***	***	-0.0622	0.4205	0.2295	0.2951	0.2924	0.8416
	Prior Day Rolling Average Trading Volume	***	***	***	-0.0037	0.088	0.9647	0.9871	0.9871	0.1137
VAR_TF	Current Volatility (RV)	***	***	***	0.0198	0.6013	-0.2683	0.2435	0.2406	0.86969
	Current Trading Volume	***	***	***	-0.0664	0.6039	0.2236	0.5538	0.5521	0.6686
	Prior Day Rolling Average Trading Volume	***	***	***	-0.0118	0.0714	0.9539	0.9938	0.9938	0.0789



Starbucks Corp. (SBUX)

Model	Response Predictors	Current Volatility	Current Trading Volume	Prior Day Rolling Average Trading Volume	Current Volatility	Current Trading Volume	Prior Day Rolling Average Trading Volume	R ²	Adjusted R ²	Residual Standard Error
VAR_TV	Current Volatility (RA)	***	***	***	0.0233	0.5303	-0.0401	0.3126	0.3100	0.8312
	Current Trading Volume	***	***	***	-0.1072	0.5394	0.1408	0.2969	0.2942	0.8406
	Prior Day Rolling Average Trading Volume	***	***	***	-0.0016	0.1027	0.9525	0.9806	0.9805	0.1385
VAR_TF	Current Volatility (RA)	***	***	***	0.0282	0.5721	-0.0446	0.3429	0.3404	0.8126
	Current Trading Volume	***	***	***	-0.1097	0.5469	0.1462	0.3140	0.3114	0.8304
	Prior Day Rolling Average Trading Volume	***	***	***	0.0107	0.0891	0.9541	0.9825	0.9824	0.1326
VAR_TV	Current Volatility (RS)	***	***	***	-0.0219	0.5081	0.0430	0.2801	0.2573	0.8623
	Current Trading Volume	***	***	***	-0.1051	0.5341	0.1397	0.2971	0.2944	0.8405
	Prior Day Rolling Average Trading Volume	***	***	***	-0.0060	0.1050	0.9527	0.9806	0.9806	0.1384
VAR_TF	Current Volatility (RS)	***	***	***	-0.0209	0.5065	0.0333	0.2942	0.2915	0.8423
	Current Trading Volume	***	***	***	-0.1085	0.5436	0.1456	0.3144	0.3118	0.8301
	Prior Day Rolling Average Trading Volume	***	***	***	0.0074	0.0912	0.9543	0.9825	0.9824	0.1327
VAR_TV	Current Volatility (RV)	***	***	***	-0.0603	0.2978	-0.0133	0.0723	0.06877	0.9604
	Current Trading Volume	***	***	***	-0.1059	0.5664	0.1236	0.3065	0.3039	0.8348
	Prior Day Rolling Average Trading Volume	***	***	***	0.0190	0.0919	0.9540	0.9809	0.9808	0.1383
VAR_TF	Current Volatility (RV)	***	***	***	-0.0694	0.3001	-0.0367	0.0917	0.0882	0.9593
	Current Trading Volume	***	***	***	-0.1796	0.5842	0.1246	0.3284	0.3268	0.8210
	Prior Day Rolling Average Trading Volume	***	***	***	0.0282	0.0781	0.9574	0.9830	0.9830	0.1305



Target Corporation (TGT)

Model	Response Predictors	Current Volatility	Current Trading Volume	Prior Day Rolling Average Trading Volume	Current Volatility	Current Trading Volume	Prior Day Rolling Average Trading Volume	R ²	Adjusted R ²	Residual Standard Error
VAR_TV	Current Volatility (RA)	***	***	***	0.1501	0.4872	-0.0231	0.3183	0.3137	0.8280
	Current Trading Volume	***	***	***	0.0522	0.3482	0.1736	0.2174	0.2144	0.8869
	Prior Day Rolling Average Trading Volume	***	***	***	0.0315	0.0755	0.9573	0.9835	0.9834	0.1288
VAR_TF	Current Volatility (RA)	***	***	***	0.0881	0.4784	-0.0727	0.2586	0.2568	0.8616
	Current Trading Volume	***	***	***	0.0042	0.5804	0.1435	0.4326	0.4305	0.7502
	Prior Day Rolling Average Trading Volume	***	***	***	0.0090	0.0916	0.9494	0.9885	0.9885	0.1074
VAR_TV	Current Volatility (RS)	***	***	***	-0.0021	0.5037	-0.0315	0.2432	0.2403	0.8700
	Current Trading Volume	***	***	***	0.0387	0.3990	0.1817	0.2163	0.2133	0.8875
	Prior Day Rolling Average Trading Volume	***	***	***	0.0225	0.0786	0.9596	0.9831	0.983	0.1303
VAR_TF	Current Volatility (RS)	***	***	***	-0.0608	0.4480	-0.0604	0.1885	0.1865	0.9009
	Current Trading Volume	***	***	***	-0.0226	0.5969	0.1400	0.4329	0.4308	0.7549
	Prior Day Rolling Average Trading Volume	***	***	***	0.0008	0.0969	0.9496	0.9885	0.9884	0.1076
VAR_TV	Current Volatility (RV)	***	***	***	0.0697	0.4822	-0.0484	0.2374	0.2345	0.8672
	Current Trading Volume	***	***	***	0.0839	0.3323	0.1852	0.2208	0.2178	0.8646
	Prior Day Rolling Average Trading Volume	***	***	***	0.0339	0.0735	0.9593	0.9836	0.9835	0.1283
VAR_TF	Current Volatility (RV)	***	***	***	0.0111	0.4505	-0.0653	0.1852	0.1821	0.8664
	Current Trading Volume	***	***	***	0.0313	0.5620	0.1480	0.4332	0.4310	0.7548
	Prior Day Rolling Average Trading Volume	***	***	***	0.0111	0.0899	0.9593	0.9885	0.9885	0.1073



Time Warner Inc. (TWX)

Model	Response Predictors	Current Volatility	Current Trading Volume	Prior Day Rolling Average Trading Volume	Current Volatility	Current Trading Volume	Prior Day Rolling Average Trading Volume	R ²	Adjusted R ²	Residual Standard Error
VAR_TV	Current Volatility (RA)	***	***	***	0.2119	0.4186	0.0025	0.2775	0.2747	0.8512
	Current Trading Volume	***	***	***	0.0927	0.3003	0.0920	0.1383	0.1351	0.9304
	Prior Day Rolling Average Trading Volume	***	***	***	0.0238	0.1012	0.9548	0.9726	0.9725	0.1658
VAR_TF	Current Volatility (RA)	***	***	***	0.1616	0.4037	-0.0231	0.2463	0.2434	0.8684
	Current Trading Volume	***	***	***	0.0578	0.4872	0.0871	0.3054	0.3038	0.8347
	Prior Day Rolling Average Trading Volume	***	***	***	0.0124	0.1003	0.9530	0.9806	0.9805	0.1386
VAR_TV	Current Volatility (RS)	***	***	***	0.1773	0.4352	0.0112	0.2624	0.2596	0.8601
	Current Trading Volume	***	***	***	0.1916	0.3004	0.0902	0.1402	0.1389	0.9294
	Prior Day Rolling Average Trading Volume	***	***	***	0.0254	0.1038	0.9548	0.9725	0.9724	0.1682
VAR_TF	Current Volatility (RS)	***	***	***	0.1386	0.4011	-0.0086	0.2220	0.2201	0.8627
	Current Trading Volume	***	***	***	0.0665	0.4883	0.0954	0.3075	0.3049	0.8340
	Prior Day Rolling Average Trading Volume	***	***	***	0.0083	0.1030	0.9627	0.9805	0.9805	0.1388
VAR_TV	Current Volatility (RV)	***	***	***	0.1533	0.3388	-0.0081	0.1603	0.1571	0.9184
	Current Trading Volume	***	***	***	0.0504	0.3151	0.0965	0.1346	0.1313	0.9325
	Prior Day Rolling Average Trading Volume	***	***	***	0.0260	0.1040	0.9559	0.9725	0.9724	0.1682
VAR_TF	Current Volatility (RV)	***	***	***	0.1279	0.2954	-0.0191	0.1290	0.1257	0.9354
	Current Trading Volume	***	***	***	0.0083	0.5128	0.0972	0.3040	0.3013	0.8362
	Prior Day Rolling Average Trading Volume	***	***	***	0.0107	0.1021	0.9532	0.9806	0.9805	0.1387

