2-2020

# Essays on the Application of the Machine Learning Methods in Finance and Policy Evaluation

Olga Guska

*The Graduate Center, City University of New York*

# Essays on the Application of the Machine Learning Methods in Finance and Policy Evaluation

by

Olga Guska

A dissertation submitted to the Graduate Faculty in Economics in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2020

This manuscript has been read and accepted by the Graduate Faculty in Economics in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

**Professor Wim Vijverberg**

| | |
|---|---|
| Date | Chair of Examining Committee |

**Professor Christos Giannikos**

| | |
|---|---|
| Date | Executive Officer |

**Professor Wim Vijverberg**

**Professor Chu-Ping Vijverberg**

**Professor Liuren Wu**

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

Essays on the Application of the Machine Learning Methods in Finance and Policy Evaluation

by

Olga Guska

Adviser: Professor Wim Vijverberg

In the age when "Big Data" is becoming almost a household word, such abundance of information in different forms and representations can be of a great help for one's decision-making let it be a trader betting on a stock, or a policy-maker assessing the potential impact of proposed regulation. Whereas traditional economic research is primarily based on the use of numerical data continuous or discrete, there is a great deal of useful information that can be extracted from text data. Such information can power novel identification strategies or help perceive solutions from a different angle, but observed volumes of such data as well as its textual format require additional preprocessing techniques. Recent expansions in the Data Science methods allow to successfully use this alternative source of information and complement traditional economic modelling with a new insight.

Abundant "Big Data" information fuels search for yet new market predictors. Some areas of research, asset pricing for instance, produced hundreds of such factors over the past few decades. However, many of these factors are weak and often lose their predicting power altogether after the date of publication. Moreover, traditional least squares approach would break down when stuffed by the hundreds of these factors, some of which are, in addition, quite highly correlated. To tame the "zoo" of the factors in a given problem, one would want to consider alternative methods suggested in the data science literature, which allow for regularization as well as potential non-linearities in the unobserved model structure.

Regularization can also be useful in building a counterfactual – a technique so often used in a policy evaluation questions. In assessment of certain legislations effect on the individual or a firm, a construction of hypothetical world in which this regulation never took place is often done through weighted averaging of the units (other individuals, or firms) which were not subject to the regulation. Calculation of weights is an optimization process stability of which requires a set of certain constraints. Some of these constraints contradict natural way of things in certain policy evaluation problems. Thus, to relax these constraints, especially in the case where the number of unaffected units is large, one could apply regularization methods and derive the weights that are suitable to the true nature of the problem.

In my dissertation, I assess the potential of usage of the machine learning methods in finance and policy evaluation. The results show that these methods prove to be useful additions to the traditional econometric approach.

This dissertation consists of two chapters.

**Chapter 1** In this study, I assess presence of herding and contrarian behavior in the stock market proxied by StockTweets authors. The personal signal of the traders about changes in the stock price is approximated by the news headlines from Reuters. I find that the market populated by the small (retail) traders which are eager to exchange their thoughts on the micro-blog is likely to exhibit presence of herding and contrarian behavior. Moreover, these social behavior estimates can be used as an insight about the stock price volatility, especially in the case of tech firms, which product/service is harder to evaluate.

**Chapter 2** The Volcker Rule (also known as Section 619 of the Dodd-Frank Act) was put into law in 2010 to regulate prop trading activities of Wall Street banks. To assess the effect of the Rule on Banks' revenues and riskiness of their activity I apply a synthetic control method in a form of the elastic net regression. The identification strategy is based on 10-K

filings text data. I find that the total gross notional amounts of derivative contracts held for the purpose other than trading would have been lower in the post-Volcker era, had there been no Volcker Rule signed into law, possibly suggesting an increase in the market-making or hedging activity of the banks. The bigger banks also show a decrease in riskiness of their activity and trading assets, although this result is statistically significant only for the few last observed months in the testing window. No significant impact of the Rule on profitability or assets available for sale has been found.

# Contents

# List of Tables

# List of Figures

# Chapter 1

# Herding in Social Media

## 1.1 Introduction

Volatility prediction is a key task in risk management: poor risk assessment might leave investors overexposed to the market fluctuations and financial institutions insufficiently capitalized. The financial meltdown of 2008 indicated a need to reassess risk-modelling practices as well as volatility forecasting tools, and to assure the validity of the forecast particularly amid most adverse economic scenarios. Up till now, the primary focus of volatility modeling was placed on the method rather than data available at hand. In this study, I consider alternative sources of information (in particular social media text data) that can be useful in the stock volatility prediction.

A good portion of stock price swings can be explained by herd behavior, a phenomenon in which traders intentionally ignore their beliefs in order to mimic someone else's action. As noted by Banerjee (1992), Bikhchandani et al. (1992), Bikhchandani and Sharma (2000), mutual mimicry among investors may temporarily move asset prices away from fundamental values, which in turn would increase the spread around the mean. Modeling herding, however, is not an easy task. In previous empirical studies, the presence of herding was analyzed through the means of statistical clustering, or it was inferred from the sequence and timing of orders to buy or sell using the generalized Glosten and Milgrom (1985) model.

Recent advances in communication and computer science have made information easily accessible to all traders. In particular, with the advent of social media, professional investor thoughts have become available to an average trader: virtual investing communities share investment recommendations and proprietary analysis; they allow for discussion, collaboration and monitoring of other participants decision making process (Oh and Sheng (2011)). People tend to pay more attention to ideas that are reinforced by conversations (Hirshleifer (2001)), and social media platforms allow such conversations to exist and thrive. Microblogging is a perfect nurturing place for ideas; sentiment emerges and spreads through the web joining the continuous flow of information that then feeds into the stock price changes. Building on this observation, recent studies show that stock microblog sentiment extracted from online traders conversations has predictive (although marginally small) power in forecasting future stock price movements (Sun et al. (2016); Oh and Sheng (2011)).

These results led me to consider herding in social media – a place where discussion between traders could be directly captured and modeled in a network context. This interaction between agents has not been considered before in empirical herding studies, and it adds to existing literature of social learning behaviors in financial markets.

Given the network structure of the StockTwits data[1], I use news-filtered trade decision in the regression that takes into account network effect. The regression coefficient captures the herding effect, which has predictive power in forecasting the stock price volatility.

[1] The StockTwits data used in this study are kindly provided by the StockTwits Inc. (stocktwits.com)

## 1.2 Herding

The stock market is organized as a continuous double auction where agents act sequentially. Such market structure allows for disproportionate impact of earlier decisions on following trades and future, sometimes long-term, outcomes. An agent might be observant of others' decisions and ignore his private information in part or in whole to follow the lead of his predecessors. The trading by an informed agent who follows the trend in past trades contradictory to his personal information is known as herd behavior. Such behavior is not irrational from a perspective that the predecessor's decisions might reflect information that the agent does not possess (asymmetry of information is imposed). In the grand scheme of things, however, imitative actions of multiple agents will not mirror their private information in full and, hence, be less informative to their successors. Banerjee (1992) finds that information reduction in the equilibrium could be quite substantial.

Unlike Banerjee (1992), Avery and Zemsky (1998) show that there are limits to price distortions caused by herding. They note that herding occurs only under a very special signal structure. In this setting, even extreme cases of herding have little effect on stock prices. The authors then investigate conditions under which herding can lead to bubbles and crashes and find such state of the world is very unlikely to happen. They show that herd behavior does not lead to excess volatility; ex ante expected market volatility is determined by fundamentals only.

Park and Sabourian (2011) reconsider these results. The authors note that the "intuitively appealing example" in Avery and Zemsky (1998) is only a special case of their own framework: they show that herding as well as contrarianism (a natural counterpart of herding) might take place in a much more likely economic setting. Moreover, both types of social learning behavior could be consistent with large movements in prices. These findings are based on the hypothesis that a trader herds only if, according to his beliefs, his private signal is generated from a fat-tailed distribution where extreme realizations are more likely than the moderate ones (a so-called U-shaped property of a signal); a trader is expected to act as a contrarian if the signal induces him to believe that it was generated by a moderate state realization (a hill-shaped signal).

The findings of the aforementioned and other theoretical studies (Bikhchandani et al. (1992), Ho Lee (1998), Cipriani and Guarino (2014) have identified mechanisms of social learning behaviors in an abstract environment. All empirical attempts to test these theories have faced the same problem: traders private information is unobservable and so it is hard to know whether the action to buy or sell was based on traders private signal or his predecessors decision. To shed some light on the problem, empiricists have tried to detect herding through clustering in trades (Lakonishok et al. (1992), Wermers (1999)), yet they themselves admit that such coordinated actions of traders might not necessarily be due to herding but be a reasonable reaction to a public announcement (Cipriani and Guarino (2014)).

Inability to test theory empirically has brought the question into the lab. There is a plethora of studies assessing herding and contrarianism in experimental setting (Cipriani and Guarino (2014); Drehmann et al. (2005)). Experimental design, while offering an advantage of complete control over selection and omitted variable bias (if conducted correctly) (J. Bloomfield et al. (2007)), has its own limitations: experiments often lack in external validity. One attempt to overcome the latter is done by Drehmann et al. (2005). Interestingly, when testing Avery and Zemsky (1998) theory in lab, they find that herd behavior does not seem to be an important force in financial markets; the agents tend to act differently from their predecessors, and their contrarian behavior, while not always being profitable, has stabilizing effect on the market as a whole. Almost a decade later, using field data, Cipriani and Guarino (2014) attempt to empirically test herding, modeled similar to Avery and Zemsky (1998). Unlike previous empirical work, the authors allow for possibility of informed traders to receive noisy signals so that a decision to ignore private information may, in fact, be optimal. In such a setting, the sequence in which orders arrive is important. The authors exploit historical transaction data for a NYSE stock to illustrate the model; they find that only 2% (4%) of the informed traders herd-buy (herd-sell) which is smaller than found in previous research using the clustering approach. They note, however, that because traders do not exploit their private information, price discovery slows down: the misalignment between market prices and fundamental values is found to be 4%.

An assumption of imperfect private signaling is gaining importance in recent literature on social learning behavior. As Drehmann et al. (2005) note, in old models with perfectly informed traders where a market-maker sets the bid-ask spread, no trade would occur; hence, the noise traders are introduced. It appears, though, that such ad hoc imposition of noise is not needed — in their experiment authors find that noise emerges automatically due to the irrationality of some trades: the distribution of irrational trades across agents is not statistically different from a distribution of trades that would have been obtained if each trader made rational decision with some probability (in their study, a probability of 0.65). The authors, however, suggest that the reason for such, to some extent irrational, decision making is not so much an extraction of relevant information but processing it correctly. In such a case, I speculate, the agents could herd not only because of asymmetry in observable information but also due to differing information processing skill sets — a novice is likely to seek advice from an experienced trader.

## 1.3 Structural model

To model the trader's behavior, I consider the following structural set-up. Let $K$ be a universe of the traders that define market. Each trader, $i, i \in [1 : K]$, makes a decision, $G_{it,s}$

to buy or sell stock, $s$, at time $t$ as a result of processing his information set, $\Omega_{it,s}$:

$$G_{it,s} = F_i(\Omega_{it,s}) \tag{1.1}$$

where the processing function, $F_i()$ is heterogeneous across the traders. The information set, $\Omega_{it,s}$, can be decomposed into a readily available, public information component about this stock, $N_{it,s}$, and an unobserved, idiosyncratic subset of the information set, $e_{it,s}$ so that

$$G_{it,s} = F_i(N_{it,s}, e_{it,s}) \tag{1.2}$$

For simplicity and interpretability of the approach, let $F_i()$ be a linear function,

$$G_{it,s} = \gamma_{0i,s} + \gamma_{1i,s} N_{t,s} + e_{it,s} \tag{1.3}$$

allowing $(\gamma_{0i,s}, \gamma_{1i,s})$ to account for the heterogeneous information-processing technology. $N_{it,s}$ is publicly available information found in the news and public reports, which could, potentially, be quantified: the natural language processing literature has grown a decent text processing tool-kit, allowing for such task to be done. However, $e_{it,s}$, is a seemingly missing piece of the puzzle, capturing anything from insider information to signals sent from other traders.

The daily trades data, a natural proxy for $G_{it,s}$ do not allow for further exploration and decomposition of the idiosyncratic subset of the information set, $e_{it,s}$. One would want to observe the interaction of those who submit the buy/sell order with other traders, $e_{-it,s}$, as well as other sources of information, $\eta_{it,s}$, so that following could be modeled:

$$e_{it,s} = H(e_{-it,s}, \eta_{it,s}) \tag{1.4}$$

The interaction between traders can be observed directly as they post their tweets on the micro-blog. Thus, making a strong assumption that the tweets reflect the traders' actions, one could obtain a partial proxy for the otherwise unobservable private signal by exploring a network connection between the trader's decisions on a given day.

Tweets, however, are collections of words, $x_{it,s}$; they are not buy or sell orders. Even so, words have meaning bearing bearish or bullish sentiment. I assume that the aggregate sentiment expressed by the collection of words is a reflection of the trader's buy/sell order. Given this logic, the following words-to-decision matching function, $\Psi$ holds:

$$G_{it,s} = \Psi(x_{it,s}|N_{it,s}, e_{it,s}) \tag{1.5}$$

## 1.4 Data

To model herding and contrarianism in social media, I use stock microblogging service (StockTwits) text data. StockTwits is a social media platform that aggregates stock-related tweets of (as of 2016) over 300,000 users, and is viewed by approximately 40 million people worldwide (Sun et al. (2016)). The content of blog is filtered to weed out finance-unrelated messages and spam, so an average participant/observer can enjoy a (relatively) high quality, large scale continuous stream of data.

The coverage of the stock in mass and social media is potentially related to the level of its liquidity. I consider stocks on the part of spectrum of their illiquidity as measured by Amihud (2002) to capture differences in the herding levels for frequently traded as well as slightly less liquid equities. The Amihud measure was calculated for all stocks traded on NYSE, NASDAQ, or AMEX exchanges during the time window covered in the study (January 2014 - December 2016) using daily absolute returns and daily dollar trading volume as obtained from CRSP and averaged over three years.[2] To assure a sufficient number of the tweet observations for each trading day, I consider the 100 most liquid stocks. Figure 1 (a) shows the distribution of log Amihud measure where left tail represents the most liquid stocks; panel (b) shows the volume of messages across the stocks collected during January 2014. Even though there is a negative relationship between the Amihud measure of illiquidity and social media coverage, the relationship is not perfectly linear: out of 100 of most liquid stocks, only 22 preserved continuity[3] in the discussion about the stock through the time window of three years. These stocks were chosen for the further analysis. The illiquidity levels as well as the stocks of the choice are presented in the Figure 2.

For each stock, from the total body of messages that cover three years (2014-2016) of microblogging history, I extract all tweets based on the presence of the stock of choice cashtag ("$AAPL" for example) in text as well as all available information about authors of these tweets.[4]. Responses to these tweets and retweets are extracted based on the initial message ID that is also saved in the StockTwits data. Such data structure allows to explore a whole dialog of StockTwits participants with regards to a specific asset.

In text pre-processing, I follow Sun et al. (2016): the text is cleaned from non-words terms, such as cashtags, numbers, URLs, and emoticons; punctuation, other symbols and unnecessary spaces also removed; all text is turned into lowercase. A resulting vector of

[2]To calculate the Amihud measure of illiquidity, all firms traded over the 2014-2016 time window were considered. Out of 8509 firms extracted from the CRSP database, only firms that were traded every day (756 days within 3 year time span) were selected. Firms with changing (multiple per PERMNO) tickers were also dropped. The illiquidity measure was constructed as $A_i = \frac{1}{D}\sum_{t=1}^{D}\frac{|r_{i,t}|}{Dvol_{i,t}\times P_{i,t}} \forall i \in [1:3741]$ where $D$ = number of days in the time window (765), $|r_{i,t}|$ = absolute daily return, $Dvol_{i,t}$ = daily number of shares traded, and $P_{i,t}$ = daily price for stock $i$. The mean Amihud illiquidity over the 5387 stocks is 0.44, which is in line with the Amihud (2002) seminal study. The Amihud illiquidity measure is reported as multiplied by $10^6$.

[3]Since the goal of this study is to calculate a daily social media herding proxy, it requires presence of such discussion (interaction) each day. Since the time span in this study is three years, it is not easy to find such firms that are constantly in a "spotlight". This is also the explanation for a slight tilt in the choice of the firms towards the Tech industry.

[4]The available data about actors includes information about the type of trader, experience, the number of followers and the number following actors and stocks, etc., however, due to the high degree of non-responsiveness to the majority of personal data questions, only followers count was used in the analysis.

messages is then trimmed by few blank messages which are an outcome of cleaning.[5] Roughly 20% of the original messages have a bullish or bearish sentiment already noted by the author, as shown in the Table 2. This part of the tweets sample is used to train the algorithm to classify all other tweets into an intention to buy or sell.

By definition, an agent herds when he ignores his private signal and follows the lead of his predecessors. To account for the private signal, I use stock-related public announcements. Reuters' "Company News" subsection in the "Stocks" section is scraped to cover the analyzed 22 stocks over 2014-2016 years.[6] Given a limited-symbol but highly informative nature of the news headlines, they are used to determine the news polarity, as described in the next Section. Duplicated news headlines are removed; the polarity is averaged across distinct news items that came out on a given day.

Once herding existence established, the question becomes whether and how it affects the stock market. As a first stab at this question, I compare the obtained daily herding estimate for each stock to the measures of the stock volatility. The realized (historical) volatility series over 10, 14, and 30 days are obtained from the OptionMetrix dataset. In addition, implied volatility surface was also extracted for the call option with expiration of 30, 60, 91, 122, 152, 182, 273, 365, 547 and 730 days over a vector of delta,[7] $\delta \in \{20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80\}$. To account for the short-term effect of herding on the stock volatility, I also construct an intra-day volatility measure for the year of 2014 following the Barndorff-Nielsen (2004) procedure and using TAQ trades data.

[5] Out of the blank messages only those without the sentiment are trimmed

[6] For example, the news story about Apple aiming for April to launch TV service with CBS, Viacom, and Starz would be located at https://www.reuters.com/finance/stocks/company-news/AAPL.O?date=10032017, the story was published on February 13, 2019 for AAPL which is traded on NASDAQ. These three inputs are used to download story as HTML data

[7] Delta of an option indicates the response in option premium to a $1.00 change in underlying price.

## 1.5 Method

The following subsections describe the steps taken in the analysis of herding in social media and its effects on the stock market. Section 5.1 describes the algorithm by which tweets of StockTweets micro-blog are classified. Section 5.2 outlines the method by which the classified tweets are related tho the public news releases. This leaves the residual component in the tweets that is subsequently entered into a network model to extract the evidence of herding and contrarian behavior, as described in Section 5.3. In Section 5.4, the herding parameters are collected and examined as a potential predictor for the stock market volatility.

### 1.5.1 Tweet classification

Regularized logit — a workhorse of machine learning (in Lee et al. (2006)) — is a simple, time-efficient yet powerful classification model which, unlike many of the machine learning methods, allows for interpretation: each covariate can be assessed with respect to how it affects the classification result. A big advantage of log-linear models in general, as compared to the discriminative machine learning models, support vector machines (SVMs) for instance, is that in addition to providing a classification rule, such models also offer an estimate of probability (Tsuruoka et al. (2009), Zhu and Hastie (2005)) and this probabilistic outcome can be further used in the text processing pipeline. It is also worth mentioning that, if executed properly, regularized logit performs on par with SVM (Zhang and Oles (2000)), which together with aforementioned benefits makes it one of the most applied models in traditional machine learning.

The training of logistic algorithm involves a maximum likelihood estimation procedure aiming to obtain the weight for the features that would maximize conditional likelihood function of the training set. In the natural language processing exercise, each feature $x_i \in \{0, 1\}$ is a binary word occurrence indicating whether the word was used in the document

$i$ (Zhang and Oles (2000)). In this application, each document is represented by the pre-processed tweet; for the purpose of classification of the remaining 80% of the tweets, the tweet-feature matrix, $x$ is created with the dimension of $n$ rows for tweets and $k$ columns for the features as observed in the training set.

Assuming that the trader's intention to buy or sell, as inferred from the tweet, is denoted by a variable $g = [1, 0]$, the probabilities through a linear function of predictors are:

$$Pr(g = 1|x) = \frac{1}{1 + e^{-(\beta_0 + x'\beta)}} \tag{1.6}$$

and

$$Pr(g = 0|x) = \frac{1}{1 + e^{(\beta_0 + x'\beta)}} = 1 - Pr(g = 1|x) \tag{1.7}$$

In text analysis, due to sparseness of $x$ and large (often larger than the number of observations) number of features extracted from the text, regularization is required. To enhance the prediction accuracy and model parsimony and interpretability, one would like to retain only the features with the strongest effect. Hense, to classify non-tagged tweets into intention to buy or sell, I use a logit algorithm with $l1$-regularization. Given the large volume of the text data, I follow Tsuruoka et al. (2009) algorithm in the execution of this classification method.

The model is fitted by regularized maximum likelihood:

$$\max_{(\beta_0,\beta) \in \mathcal{R}^{p+1}} \frac{1}{N} \sum_{i=1}^{N} \{I(g_i = 1) log p(x_i) + I(g_i = 0) log(1 - p(x_i))\} - \lambda P_\alpha(\beta) \tag{1.8}$$

where $p(x_i) = Pr(g_i = 1|x_i)$ is the probability for observation $i$ at a particular value of parameters $(\beta)$ and $P_\alpha(\beta)$ is an $l1$ penalty: $\sum_{j=1}^{k} |\beta_j|$. $\lambda$ is a hyper-parameter which controls the degree of the regularization and which is estimated via a cross-validation procedure.

Thus, one can rewrite the log-likelihood function in a following way:

$$\mathcal{L}_\beta = \sum_{i=1}^{n} L(i,\beta) - \lambda \sum_{j=1}^{k} |\beta_j| \tag{1.9}$$

where $L(i,\beta) = logp(g_i|x_i;\beta)$. To speed up the computational process, stochastic gradient descent approach is applied: a small randomly selected subset of the training set is used to approximate the gradient of the objective function in the Equation (9); the size of this subset is called a batch size and the batch size of $n$ would represent the gradient descent approach. By using smaller batch size, one can update the coefficients more frequently and accelerate the convergence. In particular, the $(p+1)^{th}$ update would be:

$$\beta^{p+1} = \beta^p + \eta_p \frac{\partial}{\partial\beta}(\sum_{i=1}^{n} L(i,\beta) - \frac{\lambda}{n} \sum_{j=1}^{k} |\beta_j|) \tag{1.10}$$

where $p$ is the iteration number and $\eta_p$ is the learning rate calculated as $\eta_p = \eta_0 \alpha^{-p/n}$ where $\alpha$ is a constant.

As Tsuruoka et al. (2009) note, the computational difficulty of the $l1$ regularization is that the last term in the Equation (10) is not differentiable when $\beta$ is zero, so the proposed solution is to update weights while dropping the regularization term, and then to adjust $\beta$'s by the difference between the absolute value of the total $l1$ penalty that each $\beta$ could have received and the total penalty that they actually received up to this point.

### 1.5.2 News polarity filtering

Recent developments in computer science and digitalization of mass media have made information more accessible to an average trader. The idea of "putting the information at one's fingertips" was championed by Bill Gates in 1980s, and over the past few decades this, at the time seemingly futuristic, idea became a reality (McFedries (2009)). The key accelerator here was the development of the so-called "small tech", in particular, small, thin client devices that would be web-connected and, hence, allow for things to be "googleable." Recent digitalization of the news reflects this trend: according to the Pew Research Center, 93% of the adults get at least some of their news online and over 90% of this number are getting their news from their mobile device (Lu (2017)).

Such a vast accessibility of news underlines my assumption that each trader receives a similar set of information each day. However, this information is not a precise signal to buy or sell. Rather, each trader processes it in a certain, heterogeneous, way and uses it in full or in part in his trade decision. Heterogeneity of private information signal processing has been suggested earlier by Drehmann et al. (2005) especially in case when information is imprecise or ambiguous.

In this study, I approximate the personal information set of each trader by the individually processed headlines of stock-related Reuters news. I assume that the public information in the form of daily news is an input into the personal signal formation process: each trader possesses a heterogeneous "technology" of how public information is translated into a private signal, and this "technology" can be captured by the parameters of the model that relates public and private information sets. The private information set is then used to execute a buy or sell decision (or in the case of StockTweets, post a tweet with a buy- or sell sentiment).

News headlines are word-count-limited, short messages that absorb the main idea from the news report. All such features of these textual data are useful to have, given limitations of the news sentiment analysis that is chosen in this study.

One of the biggest challenges of news sentiment analysis is a scarce or altogether absent training set (John and Vechtomova (2017)). When supervised learning is not feasible, the most commonly adopted approach in the identification of semantic orientation is to use polarity lexicons — the so-called bag-of-words approach. I follow the Rinkler (2019) version of such method employing lexicon specific to the financial texts Loughran and McDonald

(2011). The news polarity estimation follows Rinkler (2019) methodology, which unfolds in a following way.

Denote each news headline $s_j$ which is broken into a vector of words $\{w_1, w_2, ...w_{n_j}\}$ of the length $n$. Each word is then searched in the dictionary (lexicon) of choice. Words that were matched to positive (negative) category in the dictionary are assigned a value of $+1$ $(-1)$ and are denoted to be polarized words $(pw_i)$. Around each polarized word a polar cluster $(c_i \forall i \in [1 : k])$ is formed; each polar cluster contains polarized word itself, two immediately preceding, and two immediately succeeding words - ($w_{k,i}$, where $k \in [1 : 4]$). The words in the polar cluster are considered to be valence shifters and include:

- neutral - $(w^0_{k,i})$; have no value in equation but affect the total count of words in the headline;
- amplifier or deapmlifier - ($w^a_{k,i}$ or $w^d_{k,i}$) increase (decrease) polarity of a cluster by a multiple of 1.8.
- negator - $(w^n_{k,i})$; if the cluster contains an odd number of negators the amplifier is converted to de-amplifier and vice versa. Odd number of negators also flips the sign of the polarized word.
- adversative conjunction $(w^c_{k,i})$ also upweigths (downweigths) the cluster by 0.85 if it appears before (after) the polarized word. This is based on the belief that adversative conjunction makes the following close of a higher value while lowering the value of the previous clause.

The unbounded polarity score, $NP_j$ is then calculated as a sum of cluster values scaled by the square root of word count in the headline: $NP_j = \frac{\sum_1^k c_i}{\sqrt{n}}$. The result of this news sentiment analysis is an unbounded polarity score for each news headline.

By definition, herding (contrarianism) represents following the lead of the predecessor (trading against the predecessor) at the expense of one's own information. I assume that each actor's personal signal is created as a news release processed in a certain way for a given day. To factor in the trader's personal information set, which traders face every day, I filter obtained intention to buy or sell by the stock news polarity score, $NP_{t-1}$.[8] For each trader $i$ on day $t$ for stock $s$, this filter is modelled as follows:

$$log\frac{Pr(g_{i,t,s}=1|x_{it,s})}{Pr(g_{i,t,s}=0|x_{it,s})} = \gamma_{0i,s} + \gamma_{1i,s}NP_{t-1,s} + e_{it,s} \tag{1.11}$$

where $x_{it,s}$ denotes a set of words in which $i^{th}$ trader expresses his sentiment to buy or sell stock $s$ at the time $t$. The residual, $e_{it,s}$ is the filtered decision to buy or sell which I will subsequently use in a spatial lag regression. Note that the estimation is carried over each actor separately, delivering $(\gamma_0, \gamma_1)$ for each $i$ for each stock, which is in line with the heterogeneous information processing assumption.

In addition to the specification which uses log-odds as dependent variable in concert with the logistic regression estimated as described in the Equation (8), I consider a model in which Probability, as a continuous measure, is "filtered" by the News Polarity so that

$$Pr(g_{it,s}=1|x_{it,s}) = \alpha_{0i,s} + \alpha_{1i,s}NP_{t-1,s} + e_{it,s} \tag{1.12}$$

This is done as a robustness check for the potential imprecision in the classification executed according to the Section 4.1.

[8] To ensure that the news comes out before the tweet is posted.

### 1.5.3 Spatial (network) regression

The news-filtered decision to buy or sell captures a deviation from the actor's news-based trading rule. This part of the decision reflects trader idiosyncrasies that, in part, might be due to either irrational social behavior or unobserved subset of information that is not available to the public (insider information). In this application, I assume that the insider information is unavailable to the trader, given the nature of the proxy (StockTwits microblog) which I use for the market — such information is unlikely to be spread in social media, at least at the point when the deal is made.

To capture the commonality in the irrational part of the traders' decisions, I place these residuals, $e_{it,s}$ in a regression model that takes into account the network effect. Let $e_{st}$ be the vector of residuals $e_{it,s}$ of all recorded twits about stock $s$ on day $t$. Here, $t$ is defined as a calendar business day, for which messages related to the particular stock and responses to these messages are collected. To account for the delay in response to the issue discussed on day $t$, any message tweeted on the following day in response to the relevant messages posted on the day $t$ are also included into day $t$ and excluded from day $t+1$.

$$e_{st} = \rho_{st} W_{st} e_{st} + \epsilon_{st} \tag{1.13}$$

$$\epsilon_{st} \sim N(0, \sigma_{st}^2 I_{st}) \tag{1.14}$$

where $\rho_{st}$ is scalar parameter on the autoregressive part of spatial regression on day $t$ data. $W_{st}$ is the spatial weights matrix for the day $t$ obtained in a following way.

For ease of exposition, I omit the subscripts $s$ and $t$ for now. Let $w_{i,j}^{adj}$ be an element of spatial weights matrix, $W$, that represents the weight of influence of $j$th actor on $i$th actor's decision. Calculation of each such element is a modification of the procedure presented in Getis and Aldstadt (2004) that replaces the distance criterion by interaction and adjusts calculated weights by the significance of the affecting actor.

Step 1. Define $w^*_{i,j}$ as

$$w^*_{i,j} = \begin{cases} 1, & \text{if trader } i \text{ interacts with trader } j \\ 0, & \text{otherwise} \end{cases} \tag{1.15}$$

Here interaction is considered as a dialog between $i$th and $j$th trader rather than a unique message sent from one to another. This is done in order to avoid double-counting of traders' opinions.

Step 2. Normalize the spatial weights by the count of the followers. Customary in spatial-econometric applications, spatial weights are row-normalized: $w_{i,j} = w^*_{i,j} / \sum_{j=1}^{n} w^*_{i,j}$ such that $\sum w_{i,j} = 1$. In that case, $w_{i,j}$ is equal for each column where it is non-zero in the $i$th row, suggesting equal impact of other traders with which $i$th trader had discussions. I modify this slightly. To adjust for significance of a trader I use trader's count of followers, $m_i$:

$$w^{adj}_{i,j} = w^*_{i,j} \frac{m_j}{\sum_{j=1}^{n} m_j} \tag{1.16}$$

This procedure yields a non-symmetric direction-preserving weights matrix. In addition, I also consider a symmetric version of it, which requires a one extra step:

Step 3. Optionally, symmetricize the spatial weights matrix.

$$w^{adj,symmetric}_{i,j} = w^{adj}_{i,j} + w^{adj}_{j,i} \tag{1.17}$$

Here, the matrix would capture an interaction rather than a response, and such interaction would bear a weight of the initiator of the discussion or of the sum of initiator and the discussant if the interaction was two-sided.

Thus, the spatial weights matrix $W_{st}$ is built based on the total number of interactions between the agents (retweet, response to tweet) weighted by significance of the initiating agents which is proxied by a number of followers that they have. The matrix has a dynamic nature: new agents join StockTweets and discover each other over time, and the number of followers is also time-dependent. To capture this, I calculate the spatial weights matrix for every day. $\hat{\rho}_{t,s}$ is then obtained for each trading day for each stock; it is an estimate of the extent to which the agents on average follow their predecessors decisions, at the expense of processed public announcement. I will further refer to it as "herding rho".

Public announcements have a discrete nature and do not necessarily come out every trading day. In such cases, the private signal of a trader is a function of zero public news (or a publicly unavailable, insider information) and trades made on this day are most likely a product of information extracted from the price trend. I expect to observe an economically and statistically significant $\hat{\rho}_{t,s}$ for the trading periods without news.

It is important to point out that estimates of $\rho_{t,s}$ pertain only to herding within the micro-blog and any further generalizations are potentially hurt by a strong selection bias: the StockTweets agents represent a self-selected sample of rather small (retail) traders that post their trading decisions on the blog. Small investors typically trade much smaller volumes and are subject to higher transaction cost and less preferential treatment than institutional investors. Some research also shows that this type of trader is more prone to serious investment mistakes, especially if the trader is poorer and less educated (Campbell (2006)). Direct assessment of retail traders' behavior in stock markets faces a data problem: available surveys (i.e., SCF) do not provide high granularity data on household investments and also suffer from sampling problems. Under the assumption that micro-blog posts correctly reflect traders' decision, StockTweets data allow one to observe in detail trading strategies in a non-experimental setup. Thus, while acknowledging limitations of these data, I believe that their potential should also be recognized.

### 1.5.4 Herding and stock price volatility

The final exercise of the study is aimed at assessment of the relationship between herding and the stock price volatility. To examine this relationship, I utilize following model:

$$Vol_{t,s} = \beta^r_{0,s} + \sum_{l=1}^{10} \beta^r_{l,s}|\rho^r_{t-l,s}| + \epsilon^r_{t,s} \tag{1.18}$$

where $r$ denotes four specifications from which $\rho$ is obtained

$$r = \begin{cases} \text{directional W, log-odds as dependent} \\ \text{non-directional W, log-odds as dependent} \\ \text{directional W, probability as dependent} \\ \text{non-directional W, probability as dependent} \end{cases} \tag{1.19}$$

The construction of directional and non-directional spatial weights matrix, $W$,is explained in a previous subsection.

As a robustness check in the assessment of rho as a volatility predictor, lagged volatility is added into the right hand side of the Equation (18) so that:

$$Vol_{t,s} = \beta^r_{0,s} + \sum_{l=1}^{10} \beta^r_{l,s}|\rho^r_{t-l,s}| + \beta^r_{11,s}Vol_{t-1,s} + \epsilon^r_{t,s} \tag{1.20}$$

## 1.6 Estimation and results

### 1.6.1 Classification performance

Table 2 illustrates the proportion of of Bearish sentiment in a tagged subset of tweets. For the chosen stocks, the figure varies from the low 11% to a high 58% with the average mix of Bulls and Bears across the stocks being 2:1. This has consequences for the sentiment prediction. As Cramer (1999) notes, in a logistic analysis it is always the case that the less-frequent outcome has lower estimated prediction probabilities while more prevalent one is always predicted better. Such a dichotomy in the prediction becomes even more pronounced with a worse fit of the model.

In addition to this, the imbalance in the data also poses a problem for some of the traditional evaluation metric, such as Prediction Accuracy. Typically, in the classification exercise, the model performance is evaluated using a confusion matrix, i.e., a cross-tabulation of true and predicted class values in a testing (held-out) set. The confusion matrix includes numbers of correctly classified positive and negative (in this case, Bull and Bear) classes, which are located on the diagonal of the matrix (True Positive (TP), True negative (TN), and falsely classified positive and negative classes (False Positive (FP), False Negative (FN)), which are in the off diagonal cells. Prediction Accuracy, then, is a percentage of correctly classified positive and negative classes in all testing set. In the prediction, however, due to the imbalance, the positive class will have much higher probability to be classified correctly resulting in higher accuracy altogether.

Additional metrics, such as Precision, Recall [9], and Area under the ROC curve (AUC), are useful in properly assessing the model performance. In particular, Precision will help to detect over-predicting of positives, while from Recall one can infer which portion of true positives slipped into the negatives class. The ROC, or the Receiver Operating Characteristics, curve is a plot of all possible pairs of true positive rate (*Sensitivity* or Recall) versus false positive rate ($(1 - Specificity)$ [10]), for each value of cut-off probability measure $p \in [0, 1]$, which separates observations in a testing set into positive and negative classes; hence, the ROC curve captures the trade-off between a hit and a miss rates for the positive class at every possible probability cut-off: the curvier it is the more discriminating power the model

[9] $Precision = \frac{TP}{TP+FP}$ and $Recall = \frac{TP}{TP+FN}$

[10] False positive rate $(FPR) = \frac{FP}{TN+FP}$

has. The area under the ROC curve (AUC) is then a measure of the aggregate classification performance (Flach and Ferri (2011)).

To deal with the imbalance in the data, it is often suggested to set the probability cut-off level to the value that maximizes accuracy of classification across both classes. One of the most frequently used ways to execute such strategy is called the Youden Index (Youden (1950)). The index can be defined as

$$J = max_p\{Recall(p) + Specificity(p) - 1\} \tag{1.21}$$

where $p \in [0, 1]$. Then $J \in [0, 1]$ so that perfectly accurately classified model will receive an index of 1; 0 denotes the completely opposite case. The optimal value of the probability cut-off is then found in a following way:

$$\hat{p} = \underset{p}{\operatorname{argmax}}(J(p)) \tag{1.22}$$

where $J(p)$ is defined by the Equation (21) The frequent use of YI criterion for choosing the optimal threshold value is due to the method's simplicity and absence of requirements for additional subjective information as it is the case for the other methods (Fluss et al. (2005)).

The usage of YI probability cut-off is an example of the so-called internal approach to tackle the imbalance problem: the classification algorithm is modified to suit the data at hand (Estabrooks et al. (2004)). There is also another approach that is frequently used in the applied machine learning literature especially when the comparison of the model performance is the goal of the study: unlike the internal method, it does not use algorithm transformations that are not necessarily transferable across the different algorithms, but rather balances the training set by means of re-weighting or re-sampling. There are two ways to perform re-sampling: 1) oversampling the smaller class to make it reach the size of the larger class or 2) under-sampling the larger class to shrink it to the size of the smaller.

In this study, I use internal and external approaches to the imbalanced classes problem: I calculate Youden's Index criterion for the probability cutoff in determining which class the observation in the testing set will be assigned to; I also over-sample the smaller (usually Bear, as Table 2 would suggest) class to reach the balance in the training set and separate the classes based on the probability cut-off of 0.5. Both approaches are evaluated based on the aforementioned diagnostics measures - Accuracy, Precision, Recall, and AUC.

In the cross-validation procedure for the internal approach, the messages tagged with Bullish or Bearish sentiment are randomly split into five roughly equal groups (folds). For each fold, the estimation of the regularized logit is carried over the other four folds (combined into a training set) and probabilities prediction is executed over this fold; the procedure is repeated for each fold until a predicted probability is obtained for each observation with the sentiment tag. Then the probability cutoff is calculated via the YI procedure and the classification is completed using this cut-off. Each stock is evaluated separately due to the variation in the industry of the stock and the specifics of the events that happen in life of each stock.[11]

Similarly, for the external approach, the tagged messages are collected. The proportion, in which the Bulls and Bears are mixed, is calculated and if the set is imbalanced, the over-sampling procedure is performed for the training set. The regularized logit estimation is then carried over as described above with a small difference of omitting a probability threshold calculation. Instead the probability cut-off is set to 0.5.

The results of this exercise can be found in the Tables 3 and 4. The comparison of the diagnostic estimates suggest that over-sampling procedure fares slightly better than probability cut-off adjustment. The discrepancy is more pronounced when the imbalance is

[11] The analysis was executed over a whole body of the stocks that are less-covered in the social media and diagnostics results were compared to a stock-by-stock analysis, only marginal improvements were found. The results of this exercise are not presented in this study.

more severe, for instance, for such firms as Exxon Mobile, Starbucks, Microsoft, and General Electric. The area under the curve grows by roughly 3-4% in such cases and Accuracy and Recall also increase slightly. This, although marginal, improvement made me follow the re-sampling technique in my analysis. The classified tweets Bull and Bear mix, as predicted by the balanced training set by $l1$-regularized logit model, is depicted in Table 5.

Overall, the algorithm performance is in line with the levels suggested in the shallow natural language processing machine learning literature (Go et al. (2009)). It is important, however, to keep in mind that tweets data are non-standardized and usual pre-processing techniques might not have the same impact on the final result compared to the standard data sources such as news and financial reporting text. In addition, future research might also consider deep learning methods, which on average supply much higher accuracy outcomes (Cambria and White (2014)).

One more potential source of bias in this analysis is stemming from the sentiment tags distribution across the actors. As it was noted earlier, only a small part of the tweets have their Bullish or Bearish tag noted. In the ideal case, one would want the tagged messages to be evenly distributed across the trades to enhance accuracy of the performed classification of the leftover messages. However, it would be too optimistic to expect data to be so evenly distributed. On the other hand, extremely high concentration of the messages with sentiment among few actors would hurt classification accuracy due to the idiosyncratic usage of words across the actors. To assess the severity of tag concentration in a considered tweets set, I calculate the portion of actors who never tags for each firm separately; the portion of messages generated by these actors as well as their followers accounts are also computed and collected in Table 6. The percentage of those who systematically do not express their sentiment using Bull or Bear check box is between 26% and 60% (see Table 6) with majority of the chosen stocks having this percentage way below 50%; high concentrations of non-taggers are usually found for the stocks that are less covered by the social media stocks with a strong presence of heavily followed news re-posting actors or actors who share links to their proprietary analyses.

### 1.6.2 News impact

The news headlines have been collected for 22 firms and News Polarity was calculated according to the Rinkler (2019) method. Description statistics for the polarity time series is shown in Table 7. The average news headline has positive sentiment, however, marginally close to neutral, with the polarity score dispersed on average 0.2 units around the mean. The selection of firms presents quite varying news coverage, ranging from 96 to 669 release days within a three-year time window. The volumes of tweets and news polarity show moderately positive correlation of 0.42 suggesting that news release signal is a relevant ingredient in the trader's decision-making.

To better assess the importance of news for the average buy or sell move, I calculate daily estimated average of actors' sentiment for each stock on a given day and regress it against the lagged News Polarity variable.[12] A similar exercise was done for the daily average probability of Bullish sentiment as obtained from the regularized logistic regression; the portion of tweets with a Bullish or Bearish tag noted by the actor was assumed to have probability equal to 1 in case of Bullish sentiment and 0 in case of Bearish.

In particular, let $\overline{Y}_{i,t}$ be a value of either average daily sentiment for a day $t$, stock $i$, or average day $t$ probability of being Bullish about the stock $s$. Then, the effect of News Polarity is modeled as follows:

$$\overline{Y}_{s,t} = \beta_{0,s} + \sum_{l=1}^{10} \beta_l NP_{t-l,s} + \eta_{t,s} \tag{1.23}$$

The News polarity exemplifies stationary series with low autocorrelation levels, which

[12]News Polarity scores and average sentiment as well as average sentiment probability were tested for stationarity and according to the ADF test are found to be stationary.

allow little chance for multicolinearity issue in this set-up. (ACF plots and ADF test statistics are presented in the Table 33 and Figures 18 and 19 respectively). The coefficients on ten[13] lags of the News Polarity are presented in Tables 8, 9, 10, and 11. The results from this exercise show additional support that, especially for firms well-covered by the news the news information is relevant to the trader's decision. The significance of the coefficients vanishes almost entirely after a week of news release, suggesting a short-term nature of its impact.

This aggregate evidence supports the next step of my analysis, which is to use News Polarity to filter out the idiosyncratic news component $e_{it,s}$ for each trader. Specifically, following Equation (11), I calculate the residuals for each trader[14]; the calculations are carried out for each stock separately. To better control for possible effect of imprecisely classified decision to buy or sell, I perform a similar exercise for the continuous measure of probability, which is obtained from the regularized logit, as presented in the Equation (12).

The histograms of residuals from Equation (11) and (12) are collected and presented as the aggregates per each stock in the Figures 3, 4, 5, 6, 7, and 8. The visible skewness in the error term distributions corresponds with the imbalance in the sentiment. Figures 9, 10, and 11 present responsiveness of the actors' decisions to the freshly released news: from each regression as specified by the Equation (11) $\hat{\gamma}_{i,s}$ is collected and plotted against its P-value. The plot follows inverted V-shape with the P-value approaching 1 as the absolute value of $\hat{\gamma}_{i,s}$ converges to zero. The plot is symmetric suggesting the presence of actors with completely polar trading strategies with respect to the news. However, the number of actors who systematically uses news fresh after release in their trading decisions ($\hat{\gamma}_{i,s}$ being significant at least at the 90%-significance level, which is marked on the plot by the horizontal line) is quite low. Such a result might reflect some imprecision in the output given the simplistic

[13]Ten business days were chosen to cover two weeks. In addition, as a robustness check, weekend days were included, but the number of lags was kept the same since the predicting power of the news dies off over time closer to the end of the second week.

[14]The actors with less than 10 tweets are discarded

nature of the classification and news processing methods. However, as the previous studies state (Campbell (2006), Drehmann et al. (2005)), the average trader acts with at least some extent of irrationality, especially if such trader is a retail, non-institutional one. This gives a reason for further exploration of this irrationality and presence of herding and contrarianism in the chosen stock markets.

### 1.6.3 Herding and Contrarianism

The residuals from the Equation (11) are further used in the spatial lag regression for each stock separately, as described in the Equations (13) and (14). As a robustness check, the residuals from the Equation (12) are also utilized in the same setup. In addition, I consider two specifications of the spatial weights matrices: directional and non-directional [15] The regression is run for each day that encompasses initial tweets and any reply to the tweet that was ever posted. It is expected that there will be a lag between the initial post and a reply. However, due to the nature of micro-blogging, this lag on average does not go beyond one day, as it is shown in Table 10: the response is usually short and quick, which allows for clearer separation between the posting days.

As noted earlier, three measures of volatility are considered: 1) intraday realized volatility (calculated using Barndorff-Nielsen (2004) methodology), realized volatility calculated over 10, 14, or 30 days (as provided by the OptionMetrix data base), and implied volatility surface over the array of delta and days to expiration. Each of measures are examined in the relationship with the calculated herding $\rho$, as described by the Equation (18). The exercise is done over each of 22 considered stocks and the results of it are presented in Tables 14 - 29 and Figures 12 – 17.

Let us start with the analysis of the longer-term volatility. The performance of the four

[15]The directional spatial matrix is calculated using first two steps as described in Subsection 4.3; the non-directional matrix includes step 3.

mentioned specifications (Equation (18) and (19)) varies across the stocks, but on average the specification with the $\rho$ generated by the non-directional W matrix and log-odds as a dependent variable in the spatial regression fares slightly better than others. Table 14, 18, and 22, provide the results of this regression for the volatility calculated over 10, 14, and 30 days respectively. These tables show significant negative impact of lagged herding on volatility of major tech companies (AAPL, AMZN, NFLX, TWTR, FB), suggesting that for a longer horizon volatility herding positively impacts price discovery and brings the price closer to the fundamental value. This is true primarily for firms whose product is hard to evaluate: in this case, herding is more of a social learning type behavior allowing for rational decision of following someone else in hope that that decision is better informed than one's own belief. Collectively, the ten lagged herding rho variable explain up to 17% of variation in the volatility for some of the stocks, as noted by the adjusted R squared in the Tables. However, this number varies substantially across the stocks, being the highest for primarily tech firms, as noted earlier.

For the intraday volatility, the negative statistically significant coefficients become smaller in absolute value and insignificant (Table 26). In case of Amazon Inc. or Twitter, some lags become significant and positive, suggesting that over the short horizon (one day) herding is marginally responsible for the deviation of the stock price from its fundamental value.

As a robustness check in this exercise, I add a first lag of stock price volatility into the Equation (18) which yields the Equation (20). Adjusted R squared as well as partial R squared for the ten lags of herding Rho are presented in Tables 30 and 31.[16] For intra-day volatility, the partial R-squared for the lags of Rho reach up to 17% of unexplained variation, and in light of the adjusted R squared magnitude this suggests importance of the herding variable in the volatility modeling task. For the long-term volatility measure, the story is

[16]Two tables present results in specifications with directional and non-directional spatial matrix across longer and shorter term volatility calculated from the log-odds specification error.

not a clear cut: inclusion of the lagged volatility into the Equation (18) for this type of volatility leads to a drastic increase of adjusted R squared. The added lag absorbs virtually all (especially in the case of londer term, 30-day volatility) variation and such outcome is to be expected due to the way these variables are constructed.[17]

To account for this, similar exercise was performed on differenced longer term volatility series. The model performance dropped significantly with the adjusted R squared approaching zero; the partial R squared for the lagged herding coefficient, even though small, suggests that herding should be considered when modeling volatility. The results of this exercise are presented in the Tables 34 and 35.

For the implied volatility surface, the adjusted R squared values are collected for each delta and days to maturity combination. The resulting surface is plotted against delta and days to maturity and presented in Figures 12 and 13 for the eight (four tech and four non-tech) stocks[18]. As for the longer term volatility, the $\rho$, that is obtained from the non-directional W and log-odds as the dependent model explains more of variation in the stock volatility. The highest adjusted R squared is noted for the tech-related companies, such as Netflix, Twitter, Apple, and Amazon. For these firms, the $\rho$ generated by the non-directional W matrix and log-odds as a dependent variable in the spatial regression explains implied volatility best if days to expiration and delta are the highest.

Tables 8 – 11 show that the public information is not absorbed instantaneously but it rather takes roughly a week for the traders to process it. Taking this into account, I perform another robustness check by replacing one lag of news polarity in the Equation 11 with seven lags and recalculate herding rhos as well as their impact on volatility. The results of this robustness check are presented in the Table 32[19] and the Figures 24 – 26. In most of the

[17]For example, for 10-days volatility, the calculation of $Vol_t$ is carried over 10 past days, variation of 8 of these days will be included into the calculation of the $Vol_{t-1}$.

[18]Figures 14 – 17 show the results for the remaining firms

[19]The calculations were run only for 10 out of 22 companies with the highest volume of tweets.

cases, predictive power of such herding rhos only increases.

## 1.7 Conclusion

This study provides analysis of presence of herding and contrarian behavior in the stock market proxied by the social media micro-blog, using a novel approach of quantifying the network relationship between the traders. The study uses the simplest tools to quantify the trader's tweets with respect to their buy or sell sentiment alongside the strong assumptions that such micro-blog messages reflect actual trading behavior. Nevertheless, evidence of herding and contrarianism is found on the market of predominantly tech goods and services. Consistent with the existing literature, herding and contrarianism was found to have a negative, although marginally small, impact on the price discovery within the short time horizon, but over longer horizon social learning behavior brings the price closer to its mean, by decreasing volatility.

Table 1.1: Firms. Summary

| Firm Name | Ticker | Amihud Illiquidity | Market Cap | Sector | Industry |
|---|---|---|---|---|---|
| Apple Inc. | AAPL | 1.97E-06 | 957.81 | Technology | Computer Manufacturing |
| Amazon.com, Inc. | AMZN | 7.10E-06 | 916 | Consumer Services | Catalog/Specialty Distribution |
| Bank of America Corporation | BAC | 8.09E-06 | 289.46 | Finance | Major Banks |
| Baidu, Inc. | BIDU | 2.59E-05 | 59.5 | Technology | Computer Software: Programming, Data Processing |
| Citigroup Inc. | C | 1.08E-05 | 164.65 | Finance | Major Banks |
| Chipotle Mexican Grill, Inc. | CMG | 3.16E-05 | 19.08 | Consumer Services | Restaurants |
| eBay Inc. | EBAY | 2.64E-05 | 32.91 | Miscellaneous | Business Services |
| Ford Motor Company | F | 2.26E-05 | 37.9 | Capital Goods | Auto Manufacturing |
| Facebook, Inc. | FB | 4.46E-06 | 510.53 | Technology | Computer Software: Programming, Data Processing |
| General Electric Company | GE | 7.88E-06 | 79.43 | Energy | Consumer Electronics/Appliances |
| Gilead Sciences, Inc. | GILD | 1.15E-05 | 80.64 | Health Care | Biotechnology: Biological Products (No Diagnostic Substances) |
| General Motors Company | GM | 2.27E-05 | 58.05 | Capital Goods | Auto Manufacturing |
| Goldman Sachs Group, Inc. | GS | 1.83E-05 | 76.25 | Finance | Investment Bankers/Brokers/Service |
| International Business Machines Corporation | IBM | 1.24E-05 | 123.79 | Technology | Computer Manufacturing |
| J P Morgan Chase & Co | JPM | 9.00E-06 | 370.79 | Finance | Major Banks |
| Microsoft Corporation | MSFT | 6.08E-06 | 934.25 | Technology | Computer Software: Prepackaged Software |
| Netflix, Inc. | NFLX | 1.34E-05 | 154.88 | Consumer Services | Consumer Electronics/Video Chains |
| Procter & Gamble Company | PG | 8.87E-06 | 264.79 | Basic Industries | Package Goods/Cosmetics |
| Starbucks Corporation | SBUX | 2.11E-05 | 93.42 | Consumer Services | Restaurants |
| Tesla, Inc. | TSLA | 1.63E-05 | 47.1 | Capital Goods | Auto Manufacturing |
| Twitter, Inc. | TWTR | 3.55E-05 | 26.48 | Technology | Computer Software: Programming, Data Processing |
| Exxon Mobil Corporation | XOM | 7.49E-06 | 344.84 | Energy | Integrated Oil Companies |

The Amihud Illiquidity measure is reported as multiplied by $10^6$; Market Cap is reported in Billions of dollars. *Source: https://www.nasdaq.com/*

Table 1.2: Tweets. Summary

| Firm | Ticker | Total number of tweets | Portion of tweets with sentiment | Bearish portion of tagged tweets |
|---|---|---|---|---|
| Apple Inc. | AAPL | 1642500 | 0.246 | 0.225 |
| Amazon.com, Inc. | AMZN | 246339 | 0.205 | 0.379 |
| Bank of America Corporation | BAC | 118208 | 0.210 | 0.194 |
| Baidu, Inc. | BIDU | 45192 | 0.236 | 0.224 |
| Citigroup Inc. | C | 27739 | 0.173 | 0.324 |
| Chipotle Mexican Grill, Inc. | CMG | 78274 | 0.263 | 0.584 |
| eBay Inc. | EBAY | 23028 | 0.193 | 0.290 |
| Ford Motor Company | F | 40220 | 0.247 | 0.225 |
| Facebook, Inc. | FB | 610937 | 0.244 | 0.207 |
| General Electric Company | GE | 139374 | 0.289 | 0.133 |
| Gilead Sciences, Inc. | GILD | 216044 | 0.235 | 0.117 |
| General Motors Company | GM | 23790 | 0.174 | 0.264 |
| Goldman Sachs Group, Inc. | GS | 36350 | 0.166 | 0.332 |
| International Business Machines Corporation | IBM | 32340 | 0.197 | 0.352 |
| J P Morgan Chase & Co | JPM | 39813 | 0.168 | 0.321 |
| Microsoft Corporation | MSFT | 79043 | 0.185 | 0.175 |
| Netflix, Inc. | NFLX | 312182 | 0.264 | 0.488 |
| Procter & Gamble Company | PG | 33105 | 0.137 | 0.234 |
| Starbucks Corporation | SBUX | 45061 | 0.247 | 0.181 |
| Tesla, Inc. | TSLA | 413897 | 0.281 | 0.308 |
| Twitter, Inc. | TWTR | 485235 | 0.282 | 0.244 |
| Exxon Mobil Corporation | XOM | 121585 | 0.343 | 0.107 |

Table 1.3: Classification performance comparison. Area under the curve

| | Youden Index Approach | | Balancing |
|---|---|---|---|
| Firm | Probability cut-off | AUC | AUC |
| Apple Inc. | 0.66 | 0.756 | 0.773 |
| Amazon.com, Inc. | 0.50 | 0.779 | 0.779 |
| Bank of America Corporation | 0.69 | 0.745 | 0.758 |
| Baidu, Inc. | 0.50 | 0.768 | 0.780 |
| Citigroup Inc. | 0.50 | 0.812 | 0.812 |
| Chipotle Mexican Grill, Inc. | 0.50 | 0.769 | 0.762 |
| eBay Inc. | 0.50 | 0.766 | 0.773 |
| Ford Motor Company | 0.50 | 0.789 | 0.808 |
| Facebook, Inc. | 0.68 | 0.772 | 0.789 |
| General Electric Company | 0.69 | 0.747 | 0.789 |
| Gilead Sciences, Inc. | 0.76 | 0.741 | 0.766 |
| General Motors Company | 0.50 | 0.771 | 0.785 |
| Goldman Sachs Group, Inc. | 0.50 | 0.758 | 0.760 |
| International Business Machines Corporation | 0.50 | 0.782 | 0.786 |
| J P Morgan Chase & Co | 0.50 | 0.793 | 0.791 |
| Microsoft Corporation | 0.70 | 0.749 | 0.768 |
| Netflix, Inc. | 0.50 | 0.782 | 0.775 |
| Procter & Gamble Company | 0.50 | 0.838 | 0.850 |
| Starbucks Corporation | 0.75 | 0.740 | 0.758 |
| Tesla, Inc. | 0.59 | 0.760 | 0.770 |
| Twitter, Inc. | 0.66 | 0.767 | 0.781 |
| Exxon Mobil Corporation | 0.78 | 0.766 | 0.802 |
| Average | | 0.770 | 0.782 |

Table 1.4: Classification performance comparison. Accuracy, Precision and Recall

| Firm | Accuracy | Precision | Recall | Accuracy | Precision | Recall |
|---|---|---|---|---|---|---|
| Apple Inc. | 0.7205 | 0.8713 | 0.7501 | 0.7372 | 0.8794 | 0.7659 |
| Amazon.com, Inc. | 0.7240 | 0.7646 | 0.7989 | 0.7182 | 0.7805 | 0.7562 |
| Bank of America Corporation | 0.7078 | 0.8904 | 0.7261 | 0.7499 | 0.8886 | 0.7879 |
| Baidu, Inc. | 0.7852 | 0.8528 | 0.8741 | 0.7646 | 0.8741 | 0.8139 |
| Citigroup Inc. | 0.7502 | 0.8225 | 0.8022 | 0.7623 | 0.8293 | 0.8148 |
| Chipotle Mexican Grill, Inc. | 0.7087 | 0.6616 | 0.6131 | 0.7001 | 0.6354 | 0.6540 |
| eBay Inc. | 0.7360 | 0.8149 | 0.8104 | 0.7378 | 0.8280 | 0.7939 |
| Ford Motor Company | 0.7883 | 0.8618 | 0.8644 | 0.7844 | 0.8851 | 0.8282 |
| Facebook, Inc. | 0.7307 | 0.8887 | 0.7546 | 0.7536 | 0.8945 | 0.7810 |
| General Electric Company | 0.7554 | 0.9218 | 0.7829 | 0.7936 | 0.9322 | 0.8205 |
| Gilead Sciences, Inc. | 0.7288 | 0.9343 | 0.7449 | 0.7818 | 0.9369 | 0.8070 |
| General Motors Company | 0.7461 | 0.8313 | 0.8232 | 0.7598 | 0.8581 | 0.8083 |
| Goldman Sachs Group, Inc. | 0.7178 | 0.7837 | 0.7967 | 0.7144 | 0.7992 | 0.7636 |
| International Business Machines Corporation | 0.7294 | 0.7940 | 0.7855 | 0.7270 | 0.8031 | 0.7658 |
| J P Morgan Chase & Co | 0.7498 | 0.8070 | 0.8289 | 0.7419 | 0.8176 | 0.7966 |
| Microsoft Corporation | 0.7115 | 0.9032 | 0.7277 | 0.7733 | 0.9018 | 0.8134 |
| Netflix, Inc. | 0.7060 | 0.7123 | 0.7113 | 0.7043 | 0.6976 | 0.7428 |
| Procter & Gamble Company | 0.7965 | 0.8787 | 0.8514 | 0.8171 | 0.8981 | 0.8582 |
| Starbucks Corporation | 0.6520 | 0.9074 | 0.6395 | 0.7733 | 0.8956 | 0.8181 |
| Tesla, Inc. | 0.7075 | 0.8187 | 0.7393 | 0.7195 | 0.8270 | 0.7500 |
| Twitter, Inc. | 0.7068 | 0.8694 | 0.7198 | 0.7431 | 0.8697 | 0.7760 |
| Exxon Mobil Corporation | 0.7310 | 0.9458 | 0.7409 | 0.8151 | 0.9474 | 0.8394 |
| Average | 0.7314 | 0.8426 | 0.7675 | 0.7533 | 0.8491 | 0.7889 |

Table 1.5: Post-classification Bull and Bear sentiment. Whole sample

| Firm | Bull | Bear |
|---|---|---|
| Apple Inc. | 1107193 | 513074 |
| Amazon.com, Inc. | 151338 | 88617 |
| Bank of America Corporation | 81284 | 33663 |
| Baidu, Inc. | 16813 | 10219 |
| Citigroup Inc. | 28628 | 10705 |
| Chipotle Mexican Grill, Inc. | 427308 | 174100 |
| eBay Inc. | 100779 | 34970 |
| Ford Motor Company | 158538 | 53168 |
| Facebook, Inc. | 15821 | 7502 |
| General Electric Company | 21599 | 13811 |
| Gilead Sciences, Inc. | 20084 | 11516 |
| General Motors Company | 25959 | 13127 |
| Goldman Sachs Group, Inc. | 57527 | 19395 |
| International Business Machines Corporation | 140029 | 168766 |
| J P Morgan Chase & Co | 23839 | 8484 |
| Microsoft Corporation | 31734 | 12202 |
| Netflix, Inc. | 253712 | 155040 |
| Procter & Gamble Company | 87160 | 31032 |
| Starbucks Corporation | 31448 | 12446 |
| Tesla, Inc. | 15972 | 6430 |
| Twitter, Inc. | 35600 | 40913 |
| Exxon Mobil Corporation | 322257 | 157194 |

Table 1.6: "Bull"/"Bear" tagging

| Firm | Total actors | % of actors which do not use tag | % of tweets written by non-tagging actors | Average followers count for non-taggers | Average followers count for those who tag |
|---|---|---|---|---|---|
| Apple Inc. | 11844 | 0.264 | 0.155 | 1698 | 402 |
| Amazon.com, Inc. | 3708 | 0.345 | 0.340 | 3164 | 1138 |
| Bank of America Corporation | 1690 | 0.354 | 0.320 | 3536 | 1151 |
| Baidu, Inc. | 813 | 0.360 | 0.342 | 5597 | 1514 |
| Citigroup Inc. | 484 | 0.512 | 0.503 | 7556 | 3334 |
| Chipotle Mexican Grill, Inc. | 1369 | 0.343 | 0.324 | 4913 | 1711 |
| eBay Inc. | 437 | 0.492 | 0.521 | 7884 | 2410 |
| Ford Motor Company | 679 | 0.405 | 0.420 | 6781 | 1273 |
| Facebook, Inc. | 7300 | 0.303 | 0.244 | 1636 | 592 |
| General Electric Company | 2073 | 0.318 | 0.282 | 3322 | 526 |
| Gilead Sciences, Inc. | 2336 | 0.275 | 0.148 | 3209 | 750 |
| General Motors Company | 435 | 0.487 | 0.545 | 9075 | 2294 |
| Goldman Sachs Group, Inc. | 628 | 0.492 | 0.483 | 7862 | 4380 |
| International Business Machines Corporation | 624 | 0.489 | 0.516 | 6605 | 3138 |
| J P Morgan Chase & Co | 660 | 0.473 | 0.578 | 6851 | 3570 |
| Microsoft Corporation | 1326 | 0.428 | 0.468 | 5455 | 2320 |
| Netflix, Inc. | 4095 | 0.297 | 0.243 | 2527 | 763 |
| Procter & Gamble Company | 469 | 0.597 | 0.717 | 7932 | 3714 |
| Starbucks Corporation | 865 | 0.379 | 0.400 | 7416 | 1483 |
| Tesla, Inc. | 4920 | 0.267 | 0.199 | 2108 | 586 |
| Twitter, Inc. | 5829 | 0.263 | 0.173 | 2045 | 510 |
| Exxon Mobil Corporation | 1753 | 0.312 | 0.214 | 4336 | 765 |

Table 1.7: News coverage and news polarity

| Firm | Mean of News Polarity | SD of News Polarity | Number of days with news releases |
|---|---|---|---|
| Apple Inc. | 0.009 | 0.179 | 669 |
| Amazon.com, Inc. | 0.056 | 0.206 | 451 |
| Bank of America Corporation | 0.019 | 0.195 | 295 |
| Baidu, Inc. | 0.043 | 0.180 | 115 |
| Citigroup Inc. | 0.002 | 0.194 | 407 |
| Chipotle Mexican Grill, Inc. | 0.024 | 0.255 | 96 |
| eBay Inc. | 0.024 | 0.200 | 129 |
| Ford Motor Company | 0.020 | 0.186 | 409 |
| Facebook, Inc. | 0.007 | 0.206 | 360 |
| General Electric Company | 0.061 | 0.182 | 360 |
| Gilead Sciences, Inc. | 0.062 | 0.226 | 143 |
| General Motors Company | 0.002 | 0.174 | 526 |
| Goldman Sachs Group, Inc. | 0.018 | 0.170 | 541 |
| International Business Machines Corporation | 0.050 | 0.212 | 228 |
| J P Morgan Chase & Co | 0.000 | 0.199 | 404 |
| Microsoft Corporation | 0.020 | 0.203 | 378 |
| Netflix, Inc. | 0.027 | 0.214 | 191 |
| Procter & Gamble Company | 0.052 | 0.231 | 117 |
| Starbucks Corporation | 0.018 | 0.221 | 143 |
| Tesla, Inc. | 0.041 | 0.185 | 260 |
| Twitter, Inc. | 0.002 | 0.240 | 187 |
| Exxon Mobil Corporation | 0.012 | 0.188 | 356 |

Table 1.8: Relationship between News and Sentiment Class. Weekend Included

| | Predicted class vs News Polarity lagged by | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Firm | 1 day | 2 days | 3 days | 4 days | 5 days | 6 days | 7 days | 8 days | 9 days | 10 days |
| Apple Inc. | 0 | 0.005 | 0.003 | 0.003 | 0.007 | 0.008 | 0.014 | 0.007 | -0.001 | -0.006 |
| Amazon.com, Inc. | 0.019 | 0.016 | 0.024 | 0.013 | 0.028 | 0.035 | -0.018 | 0.007 | 0.03 | 0.043** |
| Bank of America Corporation | 0.014 | -0.004 | 0.083*** | 0.009 | -0.001 | 0.001 | -0.029 | -0.025 | -0.028 | -0.018 |
| Baidu, Inc. | 0.074 | -0.094 | -0.032 | 0.071 | -0.021 | 0.047 | 0.092 | 0.092 | 0.053 | -0.038 |
| Citigroup Inc. | 0.039 | 0.024 | -0.051 | -0.069 | -0.034 | 0.005 | 0.051 | 0.034 | -0.003 | 0.077 |
| Chipotle Mexican Grill, Inc. | 0.049 | 0.07 | -0.005 | -0.045 | 0.015 | -0.052 | 0.045 | 0.029 | 0.025 | -0.016 |
| eBay Inc. | -0.066 | -0.095 | -0.058 | 0.064 | -0.068 | -0.023 | 0.062 | -0.025 | 0.067 | -0.069 |
| Ford Motor Company | -0.017 | 0.063 | -0.091** | -0.06 | 0.031 | -0.017 | 0.049 | 0.001 | -0.031 | 0.003 |
| Facebook, Inc. | 0.003 | 0.011 | 0.036** | 0.038*** | 0.018 | 0.012 | -0.007 | -0.012 | 0.005 | 0.011 |
| General Electric Company | -0.008 | 0.033 | 0.022 | 0.015 | 0.051 | -0.021 | -0.005 | -0.044 | -0.021 | -0.003 |
| Gilead Sciences, Inc. | -0.047 | -0.005 | -0.068** | -0.032 | -0.006 | -0.029 | -0.03 | 0.019 | -0.011 | 0.024 |
| General Motors Company | 0.159*** | 0.164*** | 0.031 | 0.014 | 0.091** | 0.009 | 0.031 | 0.016 | 0.002 | 0.014 |
| Goldman Sachs Group, Inc. | 0.004 | 0.038 | -0.066 | 0.076 | -0.086** | 0.021 | 0.048 | 0.056 | -0.027 | -0.014 |
| International Business Machines Corporation | -0.019 | -0.001 | -0.01 | 0.051 | 0.053 | 0.011 | 0.072 | 0.041 | 0.045 | 0.012 |
| J P Morgan Chase & Co | -0.002 | 0.041 | 0.039 | 0.037 | 0.029 | -0.077** | -0.036 | -0.038 | -0.043 | -0.01 |
| Microsoft Corporation | 0.027 | 0.052** | -0.003 | 0.057** | -0.016 | 0.003 | 0.045 | 0.016 | -0.043 | 0.009 |
| Netflix, Inc. | 0.013 | -0.022 | 0.005 | 0.033 | 0.07*** | 0.01 | 0.005 | -0.008 | 0.009 | -0.033 |
| Procter & Gamble Company | -0.002 | 0.009 | 0.013 | 0.034 | 0.07 | -0.124** | -0.067 | -0.088 | -0.002 | -0.011 |
| Starbucks Corporation | -0.051 | 0.019 | 0.038 | -0.011 | 0.088 | -0.068 | 0.017 | -0.014 | 0.09 | -0.043 |
| Tesla, Inc. | 0.034 | 0.02 | 0.031 | 0.02 | 0.017 | -0.018 | -0.033 | -0.014 | 0.01 | -0.048 |
| Twitter, Inc. | 0.025 | 0.048 | 0.077*** | 0.087*** | 0.032 | 0.05** | 0.01 | 0.009 | 0.05** | 0.005 |
| Exxon Mobil Corporation | 0.012 | -0.006 | 0.017 | -0.066 | -0.042 | 0.039 | 0.008 | -0.034 | -0.002 | -0.073** |

Table 1.9: Relationship between News and Sentiment Probability. Weekend included

| | Predicted probability of being Bull vs News Polarity lagged by | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Firm | 1 day | 2 days | 3 days | 4 days | 5 days | 6 days | 7 days | 8 days | 9 days | 10 days |
| Apple Inc. | 0 | 0.005 | -0.001 | -0.001 | 0.006 | 0.01 | 0.01 | 0.007 | 0.003 | -0.004 |
| Amazon.com, Inc. | 0.018 | 0.018 | 0.01 | 0.014 | 0.025 | 0.024 | -0.015 | 0.001 | 0.018 | 0.036*** |
| Bank of America Corporation | 0.001 | -0.009 | 0.063*** | 0.013 | 0 | 0.003 | -0.024 | -0.01 | -0.013 | -0.011 |
| Baidu, Inc. | 0.038 | -0.087 | -0.051 | 0.03 | -0.017 | 0.026 | 0.067 | 0.045 | 0.043 | -0.013 |
| Citigroup Inc. | 0.026 | 0.021 | -0.036 | -0.051 | -0.007 | 0.027 | 0.04 | 0.026 | 0.029 | 0.055 |
| Chipotle Mexican Grill, Inc. | 0.052 | 0.076 | 0.013 | 0.006 | 0.02 | -0.033 | 0.013 | 0.031 | 0.025 | 0.025 |
| eBay Inc. | -0.094 | -0.055 | -0.068 | 0.029 | -0.097 | -0.006 | 0.022 | 0.005 | 0.032 | -0.05 |
| Ford Motor Company | -0.009 | 0.041 | -0.06** | -0.034 | 0.014 | -0.02 | 0.043 | 0.005 | -0.04 | -0.012 |
| Facebook, Inc. | 0 | 0.006 | 0.002 | -0.001 | 0.003 | -0.002 | 0.002 | -0.004 | -0.004 | 0.002 |
| General Electric Company | -0.008 | 0.023 | 0.015 | 0 | 0.031 | -0.005 | 0.004 | -0.041 | -0.017 | -0.007 |
| Gilead Sciences, Inc. | -0.021 | 0.003 | -0.054** | -0.042 | -0.027 | -0.026 | -0.03 | 0.022 | -0.028 | 0.013 |
| General Motors Company | 0.123*** | 0.122*** | 0.017 | 0.016 | 0.048 | -0.008 | 0.032 | 0.02 | -0.002 | 0.001 |
| Goldman Sachs Group, Inc. | 0.033 | 0.011 | -0.037 | 0.028 | -0.014 | 0.023 | 0.031 | 0.046 | -0.004 | -0.003 |
| International Business Machines Corporation | 0.008 | -0.018 | 0.051 | 0.043 | 0.053 | 0.026 | 0.054 | 0.046 | 0.033 | -0.018 |
| J P Morgan Chase & Co | -0.016 | 0.022 | 0.006 | 0.018 | 0.013 | -0.059** | -0.016 | -0.039 | -0.024 | -0.003 |
| Microsoft Corporation | 0.021 | 0.025 | 0.008 | 0.031 | -0.021 | 0.004 | 0.03 | 0.012 | -0.01 | 0.002 |
| Netflix, Inc. | 0.006 | -0.019 | 0.002 | 0.022 | 0.027 | 0 | 0.004 | 0.005 | 0.009 | -0.016 |
| Procter & Gamble Company | -0.006 | -0.008 | 0.02 | 0.048 | 0.019 | -0.08 | -0.013 | -0.06 | 0.032 | 0.012 |
| Starbucks Corporation | -0.017 | 0.03 | 0.026 | -0.003 | 0.026 | -0.037 | 0.005 | -0.034 | 0.056 | -0.032 |
| Tesla, Inc. | 0.038 | 0.02 | 0.018 | 0.013 | 0.017 | -0.003 | -0.027 | -0.008 | 0.008 | -0.035 |
| Twitter, Inc. | 0.024 | 0.035 | 0.068*** | 0.079*** | 0.031 | 0.053*** | 0.011 | 0.004 | 0.034 | 0.014 |
| Exxon Mobil Corporation | 0.01 | -0.004 | -0.003 | -0.041 | -0.029 | 0.039 | -0.004 | -0.042 | -0.032 | -0.057** |

Table 1.10: Relationship between News and Sentiment Class. Weekend not Included

| | Predicted class vs News Polarity lagged by | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Firm | 1 day | 2 days | 3 days | 4 days | 5 days | 6 days | 7 days | 8 days | 9 days | 10 days |
| Apple Inc. | -0.013 | 0.006 | 0.006 | 0.014 | 0.012 | -0.003 | -0.016 | -0.008 | -0.01 | -0.011 |
| Amazon.com, Inc. | 0.038** | 0.016 | 0.016 | 0.054*** | -0.002 | 0.028 | 0.015 | 0.034** | 0.015 | 0.034** |
| Bank of America Corporation | 0.011 | -0.022 | 0.014 | -0.026 | -0.025 | -0.023 | -0.012 | -0.021 | 0.006 | 0.028 |
| Baidu, Inc. | 0.023 | -0.038 | 0.049 | 0.038 | 0.045 | 0.056 | 0.011 | -0.03 | -0.043 | 0.027 |
| Citigroup Inc. | 0.023 | 0.012 | -0.027 | 0.034 | -0.026 | 0.017 | -0.002 | 0.032 | -0.047 | 0.067 |
| Chipotle Mexican Grill, Inc. | 0.066 | 0.068 | 0.044 | -0.027 | 0.034 | 0.047 | 0.034 | 0.042 | -0.011 | -0.002 |
| eBay Inc. | 0.003 | -0.046 | -0.024 | -0.028 | 0.041 | 0.007 | 0.02 | -0.061 | 0.049 | -0.077 |
| Ford Motor Company | -0.003 | 0.056 | -0.048 | -0.027 | -0.014 | -0.015 | 0.003 | 0.009 | -0.019 | 0.009 |
| Facebook, Inc. | 0.009 | 0.015 | 0.022 | 0.011 | -0.004 | -0.004 | -0.004 | 0.007 | 0.02 | -0.002 |
| General Electric Company | -0.003 | 0.025 | 0.033 | -0.018 | -0.01 | -0.008 | -0.046 | 0.011 | 0.029 | 0.044 |
| Gilead Sciences, Inc. | -0.038 | -0.021 | -0.032 | -0.037 | -0.022 | 0.026 | -0.012 | -0.014 | -0.022 | -0.018 |
| General Motors Company | 0.09*** | 0.159*** | 0.065** | 0.061 | 0.02 | -0.026 | -0.015 | 0.014 | -0.045 | 0.053 |
| Goldman Sachs Group, Inc. | 0.03 | 0.052 | -0.04 | -0.017 | 0.038 | 0.038 | 0.03 | 0.016 | -0.043 | 0.002 |
| International Business Machines Corporation | 0.013 | -0.013 | 0.045 | 0.005 | 0.092** | 0.105** | 0.014 | -0.066 | -0.05 | -0.022 |
| J P Morgan Chase & Co | 0.021 | 0.017 | 0.038 | -0.038 | -0.026 | -0.015 | -0.015 | -0.053 | 0.02 | 0.031 |
| Microsoft Corporation | 0.022 | 0.049** | 0.027 | 0.027 | 0.024 | -0.007 | -0.012 | -0.01 | -0.003 | 0 |
| Netflix, Inc. | 0.004 | -0.017 | -0.004 | 0 | -0.002 | 0 | 0.024 | -0.02 | -0.002 | -0.006 |
| Procter & Gamble Company | -0.049 | -0.003 | 0.015 | -0.058 | -0.075 | -0.065 | -0.104** | -0.039 | -0.034 | 0.001 |
| Starbucks Corporation | 0.011 | 0.039 | 0.026 | -0.007 | 0.035 | -0.03 | 0.049 | -0.019 | 0.06 | 0.042 |
| Tesla, Inc. | 0.037 | 0.022 | 0.016 | -0.02 | -0.03 | -0.029 | -0.018 | 0 | 0.014 | 0.013 |
| Twitter, Inc. | 0.027 | 0.043 | 0.064*** | 0.06*** | 0.04 | 0.029 | 0.032 | 0.038 | 0.027 | 0.045** |
| Exxon Mobil Corporation | 0.005 | -0.001 | -0.028 | -0.015 | 0.01 | -0.024 | -0.031 | -0.022 | -0.044 | 0.021 |

Table 1.11: Relationship between News and Sentiment Probability. Weekend not included

| | Predicted probability of being Bull vs News Polarity lagged by | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Firm | 1 day | 2 days | 3 days | 4 days | 5 days | 6 days | 7 days | 8 days | 9 days | 10 days |
| Apple Inc. | -0.013 | 0.005 | 0.005 | 0.013 | 0.008 | -0.001 | -0.011 | -0.006 | -0.006 | -0.009 |
| Amazon.com, Inc. | 0.028** | 0.018 | 0.013 | 0.04*** | -0.002 | 0.015 | 0.011 | 0.024 | 0.013 | 0.028** |
| Bank of America Corporation | 0.004 | -0.018 | 0.009 | -0.016 | -0.026 | -0.019 | -0.003 | -0.019 | -0.002 | 0.023 |
| Baidu, Inc. | -0.024 | -0.038 | 0.01 | 0.039 | 0.028 | 0.035 | -0.001 | -0.01 | -0.056 | 0.013 |
| Citigroup Inc. | 0.018 | 0.004 | -0.016 | 0.03 | -0.004 | 0.015 | -0.005 | 0.015 | -0.02 | 0.041 |
| Chipotle Mexican Grill, Inc. | 0.069** | 0.071** | 0.021 | -0.008 | 0.01 | 0.027 | 0.022 | 0.03 | -0.04 | 0.001 |
| eBay Inc. | -0.027 | -0.02 | -0.078 | -0.035 | -0.003 | 0.011 | -0.044 | -0.024 | -0.031 | -0.07 |
| Ford Motor Company | -0.004 | 0.043 | -0.019 | -0.024 | -0.01 | -0.005 | -0.009 | 0.002 | -0.012 | -0.021 |
| Facebook, Inc. | 0 | 0.006 | -0.002 | 0.001 | 0.001 | -0.003 | -0.002 | 0.003 | 0.004 | 0 |
| General Electric Company | -0.007 | 0.021 | 0.016 | -0.007 | -0.001 | -0.026 | -0.033 | -0.005 | 0.019 | 0.033 |
| Gilead Sciences, Inc. | -0.035 | -0.008 | -0.038 | -0.032 | -0.019 | 0.02 | -0.01 | -0.019 | -0.023 | -0.017 |
| General Motors Company | 0.08*** | 0.113*** | 0.044 | 0.037 | 0.007 | -0.017 | 0 | -0.004 | -0.024 | 0.049** |
| Goldman Sachs Group, Inc. | 0.012 | 0.018 | -0.017 | 0.007 | 0.026 | 0.034 | 0.001 | 0.022 | -0.034 | -0.009 |
| International Business Machines Corporation | 0.014 | -0.016 | 0.074** | 0.044 | 0.064** | 0.086*** | 0.026 | -0.051 | -0.042 | 0.004 |
| J P Morgan Chase & Co | 0.008 | 0.006 | 0.016 | -0.039 | -0.012 | -0.012 | -0.014 | -0.038 | 0.01 | 0.018 |
| Microsoft Corporation | 0.026 | 0.021 | 0.013 | 0.024 | 0.014 | 0.003 | 0.001 | -0.006 | -0.001 | 0.009 |
| Netflix, Inc. | 0.004 | -0.015 | -0.009 | 0 | 0 | 0.012 | 0.016 | -0.019 | -0.003 | -0.001 |
| Procter & Gamble Company | -0.025 | -0.016 | -0.012 | -0.023 | -0.012 | -0.035 | -0.071** | -0.007 | -0.03 | 0.001 |
| Starbucks Corporation | 0.018 | 0.035 | 0.007 | -0.002 | 0.02 | -0.026 | 0.034 | 0 | 0.059 | 0.005 |
| Tesla, Inc. | 0.036 | 0.019 | 0.009 | -0.01 | -0.023 | -0.024 | -0.012 | 0 | 0.008 | 0.008 |
| Twitter, Inc. | 0.026 | 0.038** | 0.061*** | 0.057*** | 0.034 | 0.031 | 0.029 | 0.031 | 0.018 | 0.037** |
| Exxon Mobil Corporation | 0.005 | -0.001 | -0.039 | -0.008 | 0.003 | -0.035 | -0.03 | -0.021 | -0.04 | 0.006 |

Table 1.12: Reply delay. Days

| Firm | Mean | SD |
|---|---|---|
| Apple Inc. | 0.08 | 0.15 |
| Amazon.com, Inc. | 0.18 | 0.48 |
| Bank of America Corporation | 0.17 | 0.68 |
| Baidu, Inc. | 0.33 | 1.10 |
| Citigroup Inc. | 0.24 | 2.30 |
| Chipotle Mexican Grill, Inc. | 1.51 | 22.30 |
| eBay Inc. | 0.33 | 0.96 |
| Ford Motor Company | 0.21 | 0.89 |
| Facebook, Inc. | 0.17 | 1.56 |
| General Electric Company | 0.26 | 0.59 |
| Gilead Sciences, Inc. | 0.10 | 0.19 |
| General Motors Company | 0.20 | 0.61 |
| Goldman Sachs Group, Inc. | 0.42 | 2.97 |
| International Business Machines Corporation | 0.37 | 2.38 |
| J P Morgan Chase & Co | 0.27 | 1.44 |
| Microsoft Corporation | 0.19 | 0.67 |
| Netflix, Inc. | 0.15 | 0.52 |
| Procter & Gamble Company | 0.40 | 2.37 |
| Starbucks Corporation | 0.19 | 0.58 |
| Tesla, Inc. | 0.11 | 0.16 |
| Twitter, Inc. | 0.10 | 0.22 |
| Exxon Mobil Corporation | 0.27 | 1.48 |

Table 1.13: Rho. Summary

| Firm | Mean | SD | Min | Max |
|---|---|---|---|---|
| Apple Inc. | 0.002 | 0.095 | -0.418 | 0.476 |
| Amazon.com, Inc. | -0.009 | 0.185 | -0.841 | 0.741 |
| Bank of America Corporation | -0.022 | 0.219 | -1.652 | 0.974 |
| Baidu, Inc. | -0.004 | 0.176 | -1.507 | 1.681 |
| Citigroup Inc. | 0.001 | 0.053 | -0.428 | 0.637 |
| Chipotle Mexican Grill, Inc. | -0.001 | 0.133 | -0.924 | 1.500 |
| eBay Inc. | 0.000 | 0.041 | -0.437 | 0.565 |
| Ford Motor Company | 0.000 | 0.161 | -1.060 | 1.239 |
| Facebook, Inc. | 0.000 | 0.163 | -0.638 | 0.843 |
| General Electric Company | -0.016 | 0.188 | -1.022 | 0.850 |
| Gilead Sciences, Inc. | -0.012 | 0.188 | -1.220 | 0.862 |
| General Motors Company | -0.002 | 0.083 | -0.910 | 0.984 |
| Goldman Sachs Group, Inc. | -0.003 | 0.084 | -0.700 | 0.733 |
| International Business Machines Corporation | -0.002 | 0.069 | -0.619 | 0.644 |
| J P Morgan Chase & Co | -0.005 | 0.101 | -1.942 | 0.528 |
| Microsoft Corporation | -0.003 | 0.136 | -1.365 | 0.815 |
| Netflix, Inc. | -0.012 | 0.170 | -1.622 | 0.700 |
| Procter & Gamble Company | 0.001 | 0.095 | -0.538 | 0.762 |
| Starbucks Corporation | -0.004 | 0.119 | -0.601 | 0.732 |
| Tesla, Inc. | -0.007 | 0.182 | -2.517 | 0.546 |
| Twitter, Inc. | -0.001 | 0.165 | -0.717 | 0.592 |
| Exxon Mobil Corporation | -0.010 | 0.213 | -3.473 | 1.180 |

Table 1.14: Realized volatility (10 days) vs Herding Rho (Non-directional W matrix, Log-odds)

| | Lags | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Firm | 1 day | 2 days | 3 days | 4 days | 5 days | 6 days | 7 days | 8 days | 9 days | 10 days | Rsq |
| AAPL | -0.176* | -0.176* | -0.188* | -0.226** | -0.237** | -0.171* | -0.187* | -0.159 | -0.089 | -0.162 | 0.04 |
| AMZN | -0.135** | -0.106* | -0.043 | -0.06 | -0.056 | -0.05 | 0.006 | 0.01 | -0.056 | -0.065 | 0.01 |
| BAC | -0.052* | -0.035 | -0.015 | -0.008 | -0.014 | -0.002 | 0.021 | 0.025 | 0.021 | 0.005 | 0.00 |
| BIDU | -0.014 | -0.004 | -0.011 | 0.001 | 0.002 | -0.014 | 0.003 | -0.012 | -0.017 | -0.021 | -0.01 |
| C | 0.01 | 0.003 | 0 | -0.011 | -0.012 | -0.013 | -0.021 | -0.014 | -0.013 | 0.006 | -0.01 |
| CMG | -0.003 | -0.031 | -0.023 | -0.023 | -0.027 | -0.033 | -0.032 | -0.037 | -0.044 | -0.031 | 0.00 |
| EBAY | 0 | -0.009 | -0.014 | -0.027 | -0.018 | -0.01 | -0.032 | -0.047 | -0.042 | -0.039 | 0.01 |
| F | 0.004 | 0.002 | 0.009 | 0.014 | 0.018 | 0.009 | 0.008 | 0.02 | 0.01 | 0.01 | -0.01 |
| FB | -0.142* | -0.124* | -0.114 | -0.106 | -0.134* | -0.133* | -0.09 | -0.125* | -0.126* | -0.098 | 0.03 |
| GE | -0.012 | -0.003 | 0.013 | 0 | 0.006 | -0.005 | -0.015 | -0.018 | -0.043** | -0.035* | 0.00 |
| GILD | -0.062 | -0.071* | -0.061 | -0.034 | -0.022 | -0.011 | 0.004 | 0.008 | 0.033 | 0.034 | 0.00 |
| GM | 0.017 | 0.005 | -0.005 | 0.005 | 0.005 | 0.006 | 0.012 | 0.001 | -0.005 | -0.023 | -0.01 |
| GS | 0.063*** | 0.057*** | 0.056*** | 0.042** | 0.025 | 0.029* | 0.033* | 0.037** | 0.04** | 0.027 | 0.11 |
| IBM | 0.016 | 0.004 | 0.012 | 0.015 | 0.019 | 0.01 | -0.003 | 0.004 | 0.002 | 0 | -0.01 |
| JPM | 0.046** | 0.046** | 0.049*** | 0.052*** | 0.04** | 0.026 | 0.012 | -0.002 | -0.015 | -0.006 | 0.04 |
| MSFT | -0.019 | -0.027 | -0.028 | -0.038 | -0.046** | -0.039* | -0.041* | -0.041* | -0.034 | -0.027 | 0.02 |
| NFLX | -0.474*** | -0.414*** | -0.393*** | -0.303*** | -0.252*** | -0.129 | -0.102 | 0.027 | 0.108 | 0.131 | 0.14 |
| PG | 0.026* | 0.028** | 0.026** | 0.018 | 0.023* | 0.023* | 0.016 | 0.014 | 0.027** | 0.016 | 0.05 |
| SBUX | 0.016 | 0.01 | 0.011 | 0.015 | 0.014 | 0.008 | 0.004 | -0.014 | -0.018 | -0.011 | -0.01 |
| TSLA | -0.216*** | -0.238*** | -0.262*** | -0.285*** | -0.255*** | -0.219*** | -0.136 | -0.017 | -0.028 | -0.027 | 0.07 |
| TWTR | -0.351** | -0.309** | -0.313** | -0.353** | -0.384*** | -0.35** | -0.233 | -0.014 | 0.027 | 0.054 | 0.04 |
| XOM | -0.015 | -0.017 | -0.023 | -0.023 | -0.023 | -0.016 | -0.019 | -0.012 | -0.009 | -0.014 | 0.00 |

Table 1.15: Realized volatility (10 days) vs Herding Rho (Directional W matrix, Log-odds

| | Lags | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Firm | 1 day | 2 days | 3 days | 4 days | 5 days | 6 days | 7 days | 8 days | 9 days | 10 days | Rsq |
| AAPL | -0.055 | -0.076 | -0.064 | -0.12* | -0.153** | -0.076 | -0.074 | -0.087 | -0.038 | -0.082 | 0.01 |
| AMZN | -0.054 | -0.036 | -0.052 | -0.047 | -0.04 | -0.053 | -0.041 | -0.04 | -0.075* | -0.067 | 0.01 |
| BAC | -0.007 | 0.007 | 0.021 | 0.01 | 0.009 | -0.003 | 0.003 | 0.009 | 0.002 | -0.017 | -0.01 |
| BIDU | 0.058* | 0.055* | 0.013 | 0.036 | 0.084*** | 0.084*** | 0.091*** | 0.089*** | 0.111*** | 0.143*** | 0.08 |
| C | -0.078 | -0.051 | -0.033 | -0.021 | 0.031 | 0.009 | -0.047 | -0.118 | -0.108 | -0.037 | -0.01 |
| CMG | 0.042 | 0.027 | 0.027 | 0.02 | 0.016 | 0.015 | 0.045 | 0.01 | 0.017 | 0.037 | -0.01 |
| EBAY | 0.063 | 0.085 | 0.195 | 0.153 | 0.193 | 0.172 | 0.094 | 0.023 | 0.037 | -0.024 | 0.00 |
| F | 0.002 | 0.005 | -0.01 | -0.012 | -0.007 | -0.021 | -0.028 | -0.042* | -0.03 | -0.027 | 0.00 |
| FB | -0.143*** | -0.129*** | -0.117** | -0.1** | -0.11** | -0.095** | -0.071 | -0.076 | -0.054 | -0.04 | 0.06 |
| GE | 0.045** | 0.05** | 0.048** | 0.048** | 0.043** | 0.04** | 0.032 | 0.034* | 0.025 | 0.022 | 0.08 |
| GILD | 0.012 | 0.007 | 0.003 | -0.019 | -0.003 | -0.012 | 0.006 | -0.006 | 0.006 | 0 | -0.01 |
| GM | -0.026 | -0.034 | -0.031 | -0.014 | -0.017 | -0.002 | -0.051 | -0.066 | -0.074* | -0.065 | 0.01 |
| GS | 0.023 | 0.012 | 0.038 | 0.043 | 0.034 | 0.023 | 0.001 | 0.025 | 0.067 | 0.064 | 0.00 |
| IBM | 0.087* | 0.073 | 0.067 | 0.01 | 0.012 | 0.012 | 0.017 | 0.055 | 0.067 | 0.075 | 0.01 |
| JPM | 0.047 | 0.04 | 0.012 | 0.041 | 0.025 | 0.024 | 0.012 | 0.044 | 0.033 | 0.01 | 0.00 |
| MSFT | 0.038 | 0.041 | 0.029 | 0.049 | 0.047 | 0.039 | 0.023 | 0.01 | 0.009 | -0.012 | 0.00 |
| NFLX | -0.136** | -0.128** | -0.15** | -0.112* | -0.079 | -0.05 | -0.021 | 0.014 | 0.023 | 0.063 | 0.01 |
| PG | 0.023 | 0.028 | 0.03 | 0.026 | 0.022 | 0.021 | 0.032 | 0.033 | 0.045* | 0.033 | 0.02 |
| SBUX | 0.033 | 0.014 | 0.006 | 0.002 | 0.013 | -0.016 | -0.019 | -0.016 | -0.006 | 0.012 | -0.01 |
| TSLA | -0.019 | -0.055 | -0.084** | -0.082* | -0.081* | -0.08* | -0.046 | -0.034 | -0.032 | -0.03 | 0.01 |
| TWTR | -0.02 | -0.013 | -0.034 | -0.05 | -0.157* | -0.073 | -0.03 | 0.059 | 0.018 | 0.041 | -0.01 |
| XOM | 0.03* | 0.04** | 0.031* | 0.028 | 0.02 | 0.027 | 0.024 | 0.028 | 0.025 | 0.018 | 0.04 |

Table 1.16: Realized volatility (10 days) vs Herding Rho (Directional W matrix, Probability

| | Lags | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Firm | 1 day | 2 days | 3 days | 4 days | 5 days | 6 days | 7 days | 8 days | 9 days | 10 days | Rsq |
| AAPL | -0.104 | -0.171** | -0.271*** | -0.316*** | -0.253*** | -0.23*** | -0.195** | -0.132* | -0.096 | -0.1 | 0.12 |
| AMZN | -0.049 | -0.031 | -0.04 | -0.06 | -0.048 | -0.041 | -0.022 | -0.033 | -0.028 | 0.025 | 0.00 |
| BAC | -0.004 | 0 | 0.011 | -0.005 | -0.012 | -0.018 | -0.006 | -0.008 | -0.004 | -0.015 | -0.01 |
| BIDU | 0.059 | 0.059 | 0.017 | 0.026 | 0.046 | 0.058 | 0.075* | 0.07* | 0.082** | 0.128*** | 0.03 |
| C | -0.038 | -0.011 | -0.021 | 0 | 0.077 | 0.066 | 0.042 | 0.001 | 0.022 | 0.039 | -0.01 |
| CMG | 0.008 | -0.001 | -0.01 | -0.001 | -0.011 | 0.008 | 0.044 | 0.014 | 0.008 | 0.049 | -0.01 |
| EBAY | 0.035 | 0.045 | 0.178 | 0.131 | 0.134 | 0.131 | 0.044 | -0.01 | 0.011 | -0.084 | -0.01 |
| F | -0.011 | -0.021 | -0.043* | -0.039 | -0.049** | -0.058** | -0.04* | -0.045* | -0.035 | -0.024 | 0.03 |
| FB | -0.237*** | -0.241*** | -0.234*** | -0.233*** | -0.205*** | -0.17*** | -0.128** | -0.093 | -0.096 | -0.12** | 0.14 |
| GE | 0.023 | 0.021 | 0.022 | 0.032 | 0.041** | 0.041** | 0.039* | 0.037* | 0.034* | 0.027 | 0.03 |
| GILD | -0.019 | -0.012 | -0.016 | -0.037 | -0.031 | -0.019 | 0.005 | 0.003 | 0.008 | 0.005 | -0.01 |
| GM | -0.001 | -0.013 | -0.024 | -0.023 | -0.039 | -0.032 | -0.077 | -0.079 | -0.086* | -0.068 | 0.01 |
| GS | 0.044 | 0.034 | 0.045 | 0.043 | 0.035 | 0.02 | 0.006 | 0.034 | 0.075* | 0.068* | 0.01 |
| IBM | 0.063 | 0.066 | 0.058 | 0.009 | 0.004 | -0.016 | -0.011 | 0.04 | 0.044 | 0.045 | 0.00 |
| JPM | 0.053 | 0.041 | 0.024 | 0.09* | 0.065 | 0.08 | 0.079 | 0.135*** | 0.121** | 0.102** | 0.02 |
| MSFT | 0.019 | 0.009 | -0.005 | 0.004 | 0.012 | 0.021 | 0.012 | 0.022 | 0.025 | 0.012 | -0.01 |
| NFLX | -0.107 | -0.066 | -0.069 | -0.068 | -0.044 | -0.027 | -0.002 | 0.046 | 0.036 | 0.027 | -0.01 |
| PG | 0.028 | 0.032 | 0.024 | 0.016 | 0.015 | 0.002 | -0.004 | 0.006 | 0.008 | 0.004 | 0.00 |
| SBUX | 0.005 | -0.009 | -0.01 | -0.008 | 0.003 | -0.009 | -0.013 | -0.008 | -0.011 | -0.004 | -0.01 |
| TSLA | 0.037 | 0 | -0.103* | -0.084 | -0.099 | -0.098 | -0.056 | -0.037 | -0.073 | -0.043 | 0.01 |
| TWTR | -0.169 | -0.082 | -0.115 | -0.072 | -0.101 | -0.053 | -0.012 | 0.13 | 0.096 | 0.117 | 0.00 |
| XOM | 0.036* | 0.039** | 0.045** | 0.038* | 0.022 | 0.023 | 0.022 | 0.028 | 0.038* | 0.032 | 0.04 |

Table 1.17: Realized volatility (10 days) vs Herding Rho (Non-directional W matrix, Probability)

| | Lags | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Firm | 1 day | 2 days | 3 days | 4 days | 5 days | 6 days | 7 days | 8 days | 9 days | 10 days | Rsq |
| AAPL | -0.216** | -0.221** | -0.277*** | -0.318*** | -0.262** | -0.238** | -0.186* | -0.118 | -0.053 | -0.059 | 0.06 |
| AMZN | -0.111* | -0.101* | -0.056 | -0.075 | -0.086 | -0.027 | 0.023 | 0.015 | -0.01 | 0.024 | 0.01 |
| BAC | -0.04 | -0.035 | -0.038 | -0.041 | -0.039 | -0.022 | 0.013 | -0.006 | 0 | -0.012 | 0.00 |
| BIDU | 0.006 | 0.006 | -0.007 | -0.013 | -0.007 | -0.025 | -0.012 | -0.026 | -0.04 | -0.029 | -0.01 |
| C | 0.017 | 0.011 | 0.011 | 0.007 | -0.005 | -0.013 | -0.018 | -0.021 | -0.019 | 0.015 | -0.01 |
| CMG | -0.02 | -0.045 | -0.031 | -0.045 | -0.02 | -0.022 | -0.03 | -0.012 | -0.009 | 0.006 | 0.00 |
| EBAY | -0.009 | -0.011 | -0.008 | -0.023 | -0.02 | -0.022 | -0.06** | -0.057** | -0.055** | -0.049* | 0.03 |
| F | -0.01 | -0.024 | -0.011 | -0.011 | 0.001 | -0.004 | 0.001 | 0.011 | 0.005 | 0.01 | -0.01 |
| FB | -0.322*** | -0.31*** | -0.341*** | -0.327*** | -0.264*** | -0.23*** | -0.167* | -0.109 | -0.123 | -0.172** | 0.13 |
| GE | -0.022 | -0.017 | 0 | 0.006 | 0.017 | 0.012 | 0.006 | -0.001 | -0.014 | -0.019 | -0.01 |
| GILD | -0.078* | -0.066 | -0.047 | -0.023 | -0.024 | -0.004 | 0.014 | 0.029 | 0.063 | 0.032 | 0.00 |
| GM | -0.002 | -0.001 | -0.016 | -0.018 | -0.015 | -0.017 | -0.003 | -0.005 | -0.006 | -0.012 | 0.00 |
| GS | 0.062*** | 0.061*** | 0.055*** | 0.048** | 0.019 | 0.022 | 0.027 | 0.031* | 0.042** | 0.024 | 0.07 |
| IBM | 0.015 | 0.008 | 0.019 | 0.014 | 0.017 | 0.007 | -0.014 | -0.001 | -0.004 | -0.011 | -0.01 |
| JPM | 0.033* | 0.032* | 0.038** | 0.042** | 0.036* | 0.025 | 0.009 | 0.001 | -0.005 | -0.003 | 0.02 |
| MSFT | -0.039 | -0.035 | -0.043* | -0.053** | -0.044* | -0.034 | -0.031 | -0.023 | -0.02 | -0.037 | 0.01 |
| NFLX | -0.436*** | -0.335*** | -0.273*** | -0.219** | -0.182** | -0.125 | -0.133 | -0.014 | 0.004 | -0.019 | 0.10 |
| PG | 0.019** | 0.02** | 0.019** | 0.015* | 0.015* | 0.012 | -0.001 | 0.004 | 0.006 | 0.001 | 0.02 |
| SBUX | 0.014 | 0.007 | 0.009 | 0.016 | 0.008 | 0.007 | 0.006 | -0.008 | -0.013 | -0.017 | -0.01 |
| TSLA | -0.153* | -0.126 | -0.25*** | -0.262*** | -0.225*** | -0.152* | -0.141* | -0.015 | -0.001 | -0.007 | 0.05 |
| TWTR | -0.444*** | -0.402*** | -0.516*** | -0.437*** | -0.368** | -0.331** | -0.368** | -0.166 | -0.155 | -0.1 | 0.08 |
| XOM | -0.017 | -0.004 | -0.005 | -0.005 | -0.007 | -0.005 | -0.004 | -0.003 | -0.003 | -0.003 | -0.01 |

Table 1.18: Realized volatility (14 days) vs Herding Rho (Non-directional W matrix, Log-odds)

| | Lags | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Firm | 1 day | 2 days | 3 days | 4 days | 5 days | 6 days | 7 days | 8 days | 9 days | 10 days | Rsq |
| AAPL | -0.138 | -0.133 | -0.155* | -0.14 | -0.179* | -0.226** | -0.257*** | -0.234** | -0.157* | -0.177* | 0.05 |
| AMZN | -0.074 | -0.053 | -0.023 | -0.089 | -0.074 | -0.071 | -0.082 | -0.055 | -0.05 | -0.051 | 0.02 |
| BAC | -0.041 | -0.03 | -0.011 | -0.011 | -0.016 | -0.009 | -0.004 | 0.003 | 0.018 | 0.024 | -0.01 |
| BIDU | -0.007 | -0.004 | 0.002 | -0.003 | -0.006 | -0.005 | -0.004 | -0.012 | -0.022 | -0.012 | -0.01 |
| C | 0.012 | 0.006 | 0 | -0.004 | -0.001 | -0.013 | -0.015 | -0.013 | -0.014 | 0 | -0.01 |
| CMG | 0.027 | 0.006 | -0.008 | -0.021 | -0.041 | -0.029 | -0.038 | -0.052* | -0.046 | -0.046 | 0.01 |
| EBAY | -0.029 | -0.034 | -0.029 | -0.026 | -0.023 | -0.008 | -0.015 | -0.037 | -0.034 | -0.036 | 0.02 |
| F | -0.001 | -0.003 | 0.007 | 0.01 | 0.008 | 0.012 | 0.008 | 0.013 | 0.012 | 0.019 | -0.01 |
| FB | -0.16** | -0.153** | -0.156** | -0.145** | -0.142** | -0.139* | -0.105 | -0.146** | -0.117 | -0.111 | 0.05 |
| GE | -0.018 | -0.022 | -0.004 | -0.01 | 0.004 | 0.007 | -0.004 | -0.012 | -0.022 | -0.028 | 0.00 |
| GILD | -0.027 | -0.047 | -0.059 | -0.04 | -0.03 | -0.036 | -0.01 | -0.003 | 0.027 | 0.034 | 0.00 |
| GM | 0.001 | 0.005 | 0.006 | 0.007 | -0.006 | 0 | 0.004 | 0.006 | 0.001 | -0.002 | -0.01 |
| GS | 0.05*** | 0.053*** | 0.05*** | 0.045*** | 0.04** | 0.038** | 0.04** | 0.037** | 0.038** | 0.035** | 0.13 |
| IBM | 0.005 | 0.005 | 0.008 | 0.003 | 0.011 | 0.011 | 0.014 | 0.012 | 0.006 | -0.004 | -0.01 |
| JPM | 0.033* | 0.034* | 0.041** | 0.042** | 0.034* | 0.032* | 0.024 | 0.017 | 0.006 | 0.009 | 0.04 |
| MSFT | -0.004 | -0.01 | -0.023 | -0.027 | -0.039* | -0.041* | -0.049** | -0.049** | -0.032 | -0.031 | 0.02 |
| NFLX | -0.384*** | -0.393*** | -0.357*** | -0.297*** | -0.27*** | -0.202** | -0.148* | -0.134* | -0.028 | 0.038 | 0.17 |
| PG | 0.019 | 0.011 | 0.017 | 0.023** | 0.03*** | 0.027** | 0.022* | 0.018 | 0.024** | 0.015 | 0.06 |
| SBUX | 0.013 | 0.004 | 0.005 | 0.02 | 0.013 | 0.003 | 0.008 | 0.004 | 0 | -0.008 | -0.01 |
| TSLA | -0.103 | -0.159** | -0.148* | -0.184** | -0.187** | -0.237*** | -0.252*** | -0.171** | -0.11 | -0.062 | 0.06 |
| TWTR | -0.258* | -0.266** | -0.296** | -0.314** | -0.344** | -0.359*** | -0.317** | -0.127 | -0.093 | 0.003 | 0.05 |
| XOM | -0.013 | -0.012 | -0.017 | -0.016 | -0.019 | -0.016 | -0.019 | -0.016 | -0.015 | -0.019 | 0.00 |

Table 1.19: .Realized volatility (14 days) vs Herding Rho (Directional W matrix, Log-odds)

| | Lags | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Firm | 1 day | 2 days | 3 days | 4 days | 5 days | 6 days | 7 days | 8 days | 9 days | 10 days | Rsq |
| AAPL | -0.037 | -0.048 | -0.032 | -0.05 | -0.104* | -0.108* | -0.125** | -0.122** | -0.077 | -0.098* | 0.01 |
| AMZN | -0.022 | -0.008 | -0.022 | -0.039 | -0.04 | -0.053 | -0.077* | -0.05 | -0.062 | -0.069* | 0.01 |
| BAC | -0.003 | 0 | 0.007 | 0.001 | 0.005 | 0.005 | -0.002 | -0.001 | 0.003 | -0.001 | -0.01 |
| BIDU | 0.027 | 0.021 | 0.014 | 0.046 | 0.069** | 0.074** | 0.095*** | 0.081*** | 0.084*** | 0.106*** | 0.08 |
| C | -0.089 | -0.079 | -0.082 | -0.064 | -0.008 | -0.034 | -0.028 | -0.031 | -0.025 | 0.01 | -0.01 |
| CMG | 0.021 | 0.009 | 0.022 | 0.024 | 0.036 | 0.04 | 0.048 | 0.029 | 0.026 | 0.035 | -0.01 |
| EBAY | 0.013 | 0.03 | 0.102 | 0.101 | 0.111 | 0.101 | 0.102 | 0.104 | 0.083 | -0.005 | 0.00 |
| F | -0.014 | -0.009 | -0.009 | -0.016 | -0.011 | -0.015 | -0.023 | -0.029 | -0.022 | -0.032 | 0.00 |
| FB | -0.143*** | -0.135*** | -0.137*** | -0.128*** | -0.108** | -0.103** | -0.085* | -0.111** | -0.073 | -0.056 | 0.06 |
| GE | 0.042** | 0.035* | 0.04** | 0.042** | 0.049** | 0.049*** | 0.044** | 0.039** | 0.032* | 0.029 | 0.08 |
| GILD | 0.007 | 0 | 0.002 | -0.006 | -0.005 | -0.02 | 0.002 | 0.006 | 0.015 | 0 | -0.01 |
| GM | -0.038 | -0.024 | -0.03 | -0.026 | -0.026 | -0.019 | -0.026 | -0.043 | -0.055 | -0.049 | 0.01 |
| GS | 0.007 | 0.003 | 0.022 | 0.027 | 0.043 | 0.033 | 0.023 | 0.03 | 0.045 | 0.05 | 0.00 |
| IBM | 0.069 | 0.072* | 0.06 | 0.036 | 0.034 | 0.023 | 0.027 | 0.048 | 0.076* | 0.06 | 0.01 |
| JPM | 0.039 | 0.024 | 0.028 | 0.049 | 0.023 | 0.024 | 0.03 | 0.032 | 0.018 | 0.024 | 0.00 |
| MSFT | 0.027 | 0.031 | 0.019 | 0.037 | 0.028 | 0.041 | 0.041 | 0.045 | 0.035 | 0.012 | 0.00 |
| NFLX | -0.098* | -0.099* | -0.112* | -0.108* | -0.096 | -0.082 | -0.047 | -0.037 | -0.014 | 0.035 | 0.01 |
| PG | 0.015 | 0.014 | 0.025 | 0.03 | 0.021 | 0.022 | 0.031 | 0.021 | 0.031 | 0.032 | 0.02 |
| SBUX | 0.006 | 0.005 | 0.004 | -0.006 | -0.009 | -0.004 | 0.012 | 0.013 | 0.014 | -0.007 | -0.01 |
| TSLA | -0.021 | -0.046 | -0.036 | -0.047 | -0.065* | -0.065* | -0.067* | -0.062 | -0.057 | -0.048 | 0.01 |
| TWTR | -0.056 | -0.043 | -0.038 | -0.034 | -0.066 | -0.042 | -0.073 | 0.009 | -0.01 | -0.008 | -0.01 |
| XOM | 0.021 | 0.026* | 0.031** | 0.04** | 0.033** | 0.036** | 0.035** | 0.021 | 0.02 | 0.02 | 0.04 |

Table 1.20: Realized volatility (14 days) vs Herding Rho (Directional W matrix, Probability

| | | | | | | Lags | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Firm | 1 day | 2 days | 3 days | 4 days | 5 days | 6 days | 7 days | 8 days | 9 days | 10 days | Rsq |
| AAPL | -0.075 | -0.1 | -0.147** | -0.179** | -0.227*** | -0.248*** | -0.239*** | -0.215*** | -0.171** | -0.163** | 0.12 |
| AMZN | -0.018 | -0.015 | -0.049 | -0.059 | -0.05 | -0.043 | -0.058 | -0.044 | -0.025 | -0.002 | 0.00 |
| BAC | -0.01 | 0.001 | 0.016 | 0 | -0.014 | -0.013 | -0.017 | -0.019 | -0.013 | -0.011 | -0.01 |
| BIDU | 0.028 | 0.026 | 0.022 | 0.046 | 0.037 | 0.046 | 0.068* | 0.056 | 0.062* | 0.09** | 0.02 |
| C | -0.038 | -0.025 | -0.041 | -0.01 | 0.051 | 0.021 | 0.036 | 0.044 | 0.056 | 0.09 | -0.01 |
| CMG | 0.017 | 0 | -0.008 | -0.007 | -0.005 | 0.017 | 0.038 | 0.02 | 0.02 | 0.023 | -0.01 |
| EBAY | 0.018 | 0.035 | 0.127 | 0.102 | 0.085 | 0.068 | 0.053 | 0.057 | 0.063 | -0.023 | -0.01 |
| F | -0.018 | -0.024 | -0.034 | -0.04* | -0.04* | -0.048** | -0.04* | -0.041* | -0.037* | -0.039* | 0.04 |
| FB | -0.198*** | -0.18*** | -0.187*** | -0.198*** | -0.208*** | -0.212*** | -0.208*** | -0.175*** | -0.151*** | -0.147*** | 0.16 |
| GE | 0.027 | 0.014 | 0.022 | 0.032 | 0.034* | 0.032 | 0.039* | 0.035* | 0.033 | 0.032 | 0.03 |
| GILD | -0.036 | -0.032 | -0.027 | -0.033 | -0.036 | -0.03 | 0 | 0.016 | 0.024 | 0.025 | -0.01 |
| GM | -0.037 | -0.024 | -0.025 | -0.028 | -0.033 | -0.028 | -0.036 | -0.044 | -0.064 | -0.057 | 0.00 |
| GS | 0.022 | 0.02 | 0.028 | 0.031 | 0.043 | 0.034 | 0.023 | 0.037 | 0.056 | 0.056 | 0.01 |
| IBM | 0.043 | 0.057 | 0.048 | 0.025 | 0.018 | -0.007 | -0.008 | 0.029 | 0.058 | 0.04 | 0.00 |
| JPM | 0.056 | 0.026 | 0.038 | 0.083* | 0.058 | 0.077 | 0.105** | 0.108** | 0.09* | 0.096** | 0.02 |
| MSFT | 0.013 | 0.016 | 0.005 | 0.014 | 0.013 | 0.026 | 0.012 | 0.03 | 0.022 | 0.01 | -0.01 |
| NFLX | -0.044 | -0.024 | -0.042 | -0.067 | -0.058 | -0.046 | -0.027 | 0.019 | 0.013 | 0.02 | -0.01 |
| PG | 0.021 | 0.016 | 0.016 | 0.023 | 0.017 | 0.013 | 0.012 | 0.004 | 0 | 0.002 | 0.00 |
| SBUX | -0.002 | -0.008 | -0.005 | -0.013 | -0.01 | -0.009 | 0 | 0.001 | 0 | -0.008 | -0.01 |
| TSLA | -0.026 | -0.034 | -0.018 | -0.009 | -0.065 | -0.071 | -0.092* | -0.067 | -0.067 | -0.061 | 0.00 |
| TWTR | -0.152 | -0.066 | -0.081 | -0.062 | -0.108 | -0.105 | -0.115 | 0.046 | 0.052 | 0.11 | 0.00 |
| XOM | 0.03* | 0.029 | 0.035* | 0.038** | 0.032* | 0.034* | 0.037** | 0.025 | 0.031* | 0.033* | 0.05 |

Table 1.21: Realized volatility (14 days) vs Herding Rho (Non-directional W matrix, Probability)

| | Lags | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Firm | 1 day | 2 days | 3 days | 4 days | 5 days | 6 days | 7 days | 8 days | 9 days | 10 days | Rsq |
| AAPL | -0.212** | -0.188** | -0.202** | -0.194** | -0.234** | -0.263*** | -0.236** | -0.18* | -0.125 | -0.13 | 0.07 |
| AMZN | -0.021 | -0.001 | -0.034 | -0.092 | -0.068 | -0.054 | -0.061 | -0.057 | -0.027 | 0.007 | 0.00 |
| BAC | -0.028 | -0.023 | -0.015 | -0.029 | -0.044 | -0.036 | -0.028 | -0.03 | -0.009 | 0.007 | 0.00 |
| BIDU | -0.011 | -0.015 | -0.002 | -0.006 | 0 | -0.011 | -0.009 | -0.024 | -0.034 | -0.029 | -0.01 |
| C | 0.018 | 0.011 | 0.003 | 0.005 | 0.004 | -0.005 | -0.003 | -0.014 | -0.017 | 0.008 | -0.01 |
| CMG | 0.017 | -0.006 | -0.023 | -0.043 | -0.03 | -0.028 | -0.027 | -0.028 | -0.021 | -0.015 | 0.00 |
| EBAY | -0.038 | -0.042 | -0.036 | -0.033 | -0.026 | -0.017 | -0.026 | -0.033 | -0.037 | -0.051* | 0.04 |
| F | -0.01 | -0.013 | -0.01 | -0.007 | -0.003 | -0.007 | -0.006 | -0.001 | 0.003 | 0.009 | -0.01 |
| FB | -0.266*** | -0.24*** | -0.248*** | -0.261*** | -0.275*** | -0.282*** | -0.271*** | -0.224*** | -0.199** | -0.211** | 0.15 |
| GE | -0.025 | -0.028 | -0.012 | -0.008 | 0.008 | 0.003 | 0.006 | 0 | 0.003 | -0.005 | -0.01 |
| GILD | -0.073** | -0.064* | -0.06* | -0.022 | -0.024 | -0.011 | 0.006 | 0.006 | 0.037 | 0.059* | 0.01 |
| GM | -0.008 | 0 | -0.008 | -0.01 | -0.02 | -0.022 | -0.01 | -0.007 | -0.005 | -0.005 | 0.00 |
| GS | 0.055*** | 0.055*** | 0.057*** | 0.045*** | 0.039** | 0.034* | 0.035** | 0.03* | 0.036** | 0.032* | 0.08 |
| IBM | 0.018 | 0.018 | 0.015 | 0.004 | 0.005 | 0.001 | 0.001 | 0.007 | 0.001 | -0.009 | -0.01 |
| JPM | 0.022 | 0.019 | 0.028 | 0.032* | 0.03* | 0.029* | 0.023 | 0.017 | 0.009 | 0.003 | 0.02 |
| MSFT | -0.02 | -0.018 | -0.029 | -0.036 | -0.038* | -0.04* | -0.047** | -0.048** | -0.037 | -0.04* | 0.02 |
| NFLX | -0.353*** | -0.33*** | -0.236*** | -0.236*** | -0.193** | -0.207** | -0.192** | -0.12 | -0.073 | -0.053 | 0.12 |
| PG | 0.014* | 0.011 | 0.012 | 0.013* | 0.014* | 0.012 | 0.011 | 0.011 | 0.012 | 0.007 | 0.03 |
| SBUX | 0.005 | 0.004 | 0.007 | 0.014 | 0.005 | 0.001 | 0.008 | 0.006 | -0.001 | -0.012 | -0.01 |
| TSLA | -0.124* | -0.101 | -0.104 | -0.129* | -0.164** | -0.182** | -0.232*** | -0.146** | -0.088 | -0.057 | 0.05 |
| TWTR | -0.408*** | -0.358*** | -0.411*** | -0.389*** | -0.376*** | -0.367*** | -0.474*** | -0.315** | -0.246* | -0.103 | 0.10 |
| XOM | -0.009 | -0.003 | -0.004 | -0.004 | -0.004 | -0.003 | -0.004 | -0.004 | -0.004 | -0.004 | -0.01 |

Table 1.22: Realized volatility (30 days) vs Herding Rho (Non-directional W matrix, Log-odds)

| | Lags | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Firm | 1 day | 2 days | 3 days | 4 days | 5 days | 6 days | 7 days | 8 days | 9 days | 10 days | Rsq |
| AAPL | -0.114 | -0.11 | -0.114 | -0.107 | -0.116 | -0.099 | -0.136* | -0.15** | -0.129* | -0.147** | 0.04 |
| AMZN | -0.017 | -0.026 | -0.028 | -0.024 | -0.018 | -0.038 | -0.047 | -0.032 | -0.065 | -0.095** | 0.01 |
| BAC | -0.002 | -0.004 | -0.001 | -0.009 | -0.005 | -0.006 | -0.003 | -0.001 | 0.004 | 0.003 | -0.01 |
| BIDU | 0.013 | 0.013 | 0.014 | 0.013 | 0.014 | 0.007 | 0.005 | -0.004 | -0.007 | -0.009 | -0.01 |
| C | 0 | 0.001 | -0.001 | -0.006 | -0.006 | -0.008 | -0.008 | -0.005 | -0.006 | -0.001 | -0.01 |
| CMG | 0.029 | 0.02 | 0.017 | 0.019 | 0.01 | 0.013 | 0.011 | 0.009 | -0.007 | -0.006 | -0.01 |
| EBAY | -0.032 | -0.037* | -0.038* | -0.045** | -0.045** | -0.039* | -0.042** | -0.043** | -0.038* | -0.033* | 0.08 |
| F | -0.012 | -0.013 | -0.009 | -0.005 | -0.003 | 0 | 0 | 0.001 | 0.001 | 0.007 | -0.01 |
| FB | -0.135** | -0.145** | -0.16*** | -0.16*** | -0.16*** | -0.162*** | -0.141** | -0.161*** | -0.145** | -0.148** | 0.09 |
| GE | -0.014 | -0.015 | -0.008 | -0.013 | -0.009 | -0.012 | -0.015 | -0.015 | -0.019 | -0.021 | 0.01 |
| GILD | -0.002 | -0.009 | -0.017 | -0.015 | -0.003 | -0.015 | -0.008 | -0.012 | -0.011 | -0.012 | -0.01 |
| GM | 0.009 | 0.009 | 0.01 | 0.007 | 0.003 | 0.005 | 0.006 | 0.001 | 0.002 | -0.001 | -0.01 |
| GS | 0.02 | 0.02 | 0.022 | 0.024* | 0.019 | 0.022 | 0.028** | 0.03** | 0.035*** | 0.036*** | 0.06 |
| IBM | 0.018 | 0.009 | 0.009 | 0.007 | 0.01 | 0.013 | 0.008 | 0.01 | 0.003 | -0.003 | 0.00 |
| JPM | 0.02 | 0.02 | 0.022 | 0.02 | 0.019 | 0.022 | 0.017 | 0.017 | 0.017 | 0.021 | 0.02 |
| MSFT | 0.013 | 0.009 | 0.011 | 0.004 | -0.009 | -0.005 | -0.016 | -0.019 | -0.013 | -0.013 | -0.01 |
| NFLX | -0.222*** | -0.245*** | -0.202*** | -0.169*** | -0.156** | -0.147** | -0.156** | -0.194*** | -0.16** | -0.146** | 0.16 |
| PG | 0.02** | 0.017* | 0.017* | 0.017* | 0.019** | 0.018** | 0.018** | 0.017* | 0.022** | 0.019** | 0.08 |
| SBUX | 0.012 | 0.007 | 0.008 | 0.012 | 0.011 | 0.008 | 0.008 | 0.003 | 0.001 | -0.001 | -0.01 |
| TSLA | -0.048 | -0.052 | -0.059 | -0.061 | -0.054 | -0.062 | -0.071 | -0.065 | -0.061 | -0.073 | 0.00 |
| TWTR | -0.054 | -0.052 | -0.053 | -0.096 | -0.115 | -0.108 | -0.191* | -0.123 | -0.122 | -0.125 | 0.00 |
| XOM | -0.016 | -0.017 | -0.018 | -0.017 | -0.015 | -0.013 | -0.017 | -0.015 | -0.016 | -0.018 | 0.01 |

Table 1.23: Realized volatility (30 days) vs Herding Rho (Directional W matrix, Log-odds)

| | Lags | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Firm | 1 day | 2 days | 3 days | 4 days | 5 days | 6 days | 7 days | 8 days | 9 days | 10 days | Rsq |
| AAPL | -0.041 | -0.043 | -0.032 | -0.039 | -0.045 | -0.027 | -0.05 | -0.055 | -0.052 | -0.07 | 0.00 |
| AMZN | -0.017 | -0.018 | -0.027 | -0.029 | -0.02 | -0.032 | -0.041 | -0.036 | -0.033 | -0.043 | 0.00 |
| BAC | -0.017 | -0.017 | -0.014 | -0.017 | -0.014 | -0.014 | -0.011 | -0.005 | -0.002 | -0.001 | 0.00 |
| BIDU | -0.084 | -0.084 | -0.077 | -0.074 | -0.057 | -0.068 | -0.069 | -0.077 | -0.082 | -0.046 | 0.00 |
| C | -0.021 | -0.021 | -0.024 | -0.029* | -0.022 | -0.026* | -0.035** | -0.032** | -0.027* | -0.028* | 0.04 |
| CMG | -0.076* | -0.082** | -0.098** | -0.101*** | -0.094** | -0.099** | -0.094** | -0.104*** | -0.09** | -0.096** | 0.09 |
| EBAY | 0.041*** | 0.036** | 0.039** | 0.038** | 0.039** | 0.037** | 0.036** | 0.035** | 0.033** | 0.037** | 0.11 |
| F | -0.001 | 0.004 | 0.011 | -0.001 | 0.004 | -0.001 | 0 | -0.001 | 0.001 | -0.009 | -0.01 |
| FB | -0.038 | -0.028 | -0.028 | -0.039 | -0.044* | -0.038 | -0.041 | -0.041 | -0.046* | -0.04 | 0.04 |
| GE | -0.005 | -0.01 | -0.002 | 0.005 | 0.015 | 0.006 | 0.013 | 0.018 | 0.032 | 0.029 | -0.01 |
| GILD | 0.02 | 0.01 | 0.01 | 0.008 | 0.012 | 0.02 | 0.03 | 0.054 | 0.051 | 0.053 | 0.00 |
| GM | 0.034 | 0.033 | 0.034 | 0.033 | 0.034 | 0.035 | 0.032 | 0.041 | 0.04 | 0.034 | 0.01 |
| GS | 0.021 | 0.017 | 0.02 | 0.029 | 0.027 | 0.029 | 0.029 | 0.03 | 0.016 | 0.011 | 0.00 |
| IBM | -0.115** | -0.139*** | -0.136*** | -0.11** | -0.093** | -0.092** | -0.078* | -0.057 | -0.033 | -0.002 | 0.04 |
| JPM | 0.017 | 0.011 | 0.014 | 0.015 | 0.016 | 0.016 | 0.024 | 0.02 | 0.024 | 0.028* | 0.01 |
| MSFT | 0.017 | 0.011 | 0.009 | 0.003 | -0.001 | 0.003 | -0.001 | -0.007 | -0.008 | -0.004 | -0.01 |
| NFLX | -0.003 | -0.01 | -0.011 | -0.018 | -0.022 | -0.023 | -0.019 | -0.024 | -0.023 | -0.033 | -0.01 |
| PG | 0.013 | 0.016 | 0.017 | 0.021 | 0.02 | 0.026* | 0.026* | 0.026* | 0.03** | 0.026* | 0.04 |
| SBUX | 0.029 | 0.024 | 0.021 | 0.035 | 0.056** | 0.057** | 0.058** | 0.058** | 0.059** | 0.065*** | 0.05 |
| TSLA | -0.016 | -0.021 | 0.014 | 0.018 | 0.033 | 0.023 | 0.011 | 0.013 | 0.009 | -0.022 | -0.01 |
| TWTR | 0.087*** | 0.076** | 0.062** | 0.056* | 0.052* | 0.057* | 0.049 | 0.041 | 0.022 | 0.023 | 0.04 |
| XOM | 0.028 | 0.001 | -0.009 | -0.033 | -0.041 | -0.031 | -0.027 | 0.007 | -0.019 | -0.011 | -0.01 |

Table 1.24: Realized volatility (30 days) vs Herding Rho (Directional W matrix, Probability)

| | Lags | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Firm | 1 day | 2 days | 3 days | 4 days | 5 days | 6 days | 7 days | 8 days | 9 days | 10 days | Rsq |
| AAPL | -0.039 | -0.032 | -0.059 | -0.081 | -0.091 | -0.098 | -0.098 | -0.102* | -0.104* | -0.107* | 0.03 |
| AMZN | -0.026 | -0.018 | -0.021 | -0.028 | -0.018 | -0.012 | -0.015 | -0.007 | 0.01 | 0.006 | -0.01 |
| BAC | -0.021 | -0.019 | -0.012 | -0.018 | -0.018 | -0.015 | -0.01 | -0.013 | -0.01 | -0.008 | 0.00 |
| BIDU | 0.028 | 0.022 | 0.02 | 0.026 | 0.036 | 0.036 | 0.042 | 0.045 | 0.047 | 0.062** | 0.01 |
| C | -0.045 | -0.039 | -0.037 | -0.035 | -0.01 | -0.026 | -0.019 | -0.02 | -0.018 | 0.017 | -0.01 |
| CMG | 0.064** | 0.053* | 0.03 | 0.033 | 0.021 | 0.034 | 0.036 | 0.024 | 0.019 | 0.024 | 0.01 |
| EBAY | -0.001 | 0.002 | 0.043 | 0.044 | 0.049 | 0.046 | 0.045 | 0.044 | 0.05 | 0 | -0.01 |
| F | -0.021 | -0.026 | -0.034** | -0.04** | -0.036** | -0.04** | -0.044*** | -0.044*** | -0.038** | -0.036** | 0.07 |
| FB | -0.145*** | -0.141*** | -0.149*** | -0.159*** | -0.15*** | -0.157*** | -0.155*** | -0.148*** | -0.159*** | -0.173*** | 0.15 |
| GE | 0.026 | 0.025 | 0.027 | 0.032* | 0.034** | 0.034** | 0.035** | 0.033** | 0.032* | 0.034** | 0.05 |
| GILD | -0.04 | -0.036 | -0.028 | -0.037 | -0.032 | -0.027 | -0.035 | -0.03 | -0.023 | -0.022 | 0.01 |
| GM | -0.022 | -0.007 | -0.015 | -0.026 | -0.032 | -0.03 | -0.035 | -0.042 | -0.056* | -0.047 | 0.01 |
| GS | 0.021 | 0.014 | 0.011 | 0.01 | 0.024 | 0.015 | 0.015 | 0.02 | 0.036 | 0.034 | 0.00 |
| IBM | -0.002 | -0.01 | -0.009 | -0.014 | -0.011 | -0.006 | 0.008 | 0.035 | 0.044 | 0.044 | -0.01 |
| JPM | 0.03 | 0.031 | 0.032 | 0.039 | 0.044 | 0.059 | 0.062 | 0.079** | 0.081** | 0.075* | 0.01 |
| MSFT | 0.01 | 0.008 | 0.011 | 0.017 | 0.014 | 0.017 | 0.014 | 0.019 | 0.009 | 0.007 | -0.01 |
| NFLX | -0.009 | -0.026 | -0.053 | -0.043 | -0.027 | -0.03 | -0.041 | 0.002 | 0.004 | 0.025 | -0.01 |
| PG | 0.027* | 0.023* | 0.017 | 0.02 | 0.019 | 0.015 | 0.015 | 0.011 | 0.013 | 0.011 | 0.02 |
| SBUX | 0.006 | 0.003 | 0.001 | -0.005 | -0.005 | -0.004 | -0.003 | -0.004 | -0.005 | -0.005 | -0.01 |
| TSLA | 0.013 | 0.005 | -0.019 | -0.016 | -0.021 | -0.044 | -0.04 | -0.024 | -0.026 | -0.037 | -0.01 |
| TWTR | -0.023 | -0.047 | -0.071 | -0.076 | -0.078 | -0.083 | -0.076 | -0.071 | -0.054 | -0.014 | -0.01 |
| XOM | 0.023 | 0.022 | 0.022 | 0.022 | 0.02 | 0.024 | 0.026* | 0.027* | 0.033** | 0.032** | 0.04 |

Table 1.25: Realized volatility (30 days) vs Herding Rho (Non-directional W matrix, Probability)

| | Lags | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Firm | 1 day | 2 days | 3 days | 4 days | 5 days | 6 days | 7 days | 8 days | 9 days | 10 days | Rsq |
| AAPL | -0.127* | -0.109 | -0.113 | -0.125* | -0.133* | -0.115 | -0.116 | -0.113 | -0.091 | -0.105 | 0.03 |
| AMZN | -0.048 | -0.032 | -0.027 | -0.035 | -0.03 | -0.012 | -0.014 | 0 | -0.006 | -0.024 | -0.01 |
| BAC | 0.006 | 0.002 | 0.003 | -0.01 | -0.014 | -0.009 | -0.006 | -0.015 | -0.016 | -0.016 | -0.01 |
| BIDU | 0.014 | 0.009 | 0.006 | 0.004 | 0.004 | -0.004 | -0.003 | -0.01 | -0.017 | -0.021 | -0.01 |
| C | 0.005 | 0.006 | 0.003 | 0 | -0.001 | -0.005 | -0.004 | -0.006 | -0.009 | 0.003 | -0.01 |
| CMG | 0.019 | 0.006 | 0.007 | 0.006 | 0.009 | 0.016 | 0.012 | 0.014 | 0.019 | 0.017 | -0.01 |
| EBAY | -0.025 | -0.035* | -0.04** | -0.047** | -0.045** | -0.045** | -0.053*** | -0.05** | -0.049** | -0.047** | 0.12 |
| F | -0.011 | -0.014 | -0.008 | -0.005 | 0 | 0.001 | -0.001 | -0.002 | -0.002 | -0.002 | -0.01 |
| FB | -0.151** | -0.172** | -0.18** | -0.191*** | -0.191*** | -0.199*** | -0.202*** | -0.192*** | -0.214*** | -0.234*** | 0.12 |
| GE | -0.02 | -0.022 | -0.019 | -0.017 | -0.011 | -0.014 | -0.013 | -0.018 | -0.017 | -0.016 | 0.01 |
| GILD | -0.028 | -0.033 | -0.035 | -0.024 | -0.012 | -0.023 | -0.019 | -0.021 | -0.012 | -0.005 | 0.01 |
| GM | 0.002 | 0.004 | 0.003 | -0.001 | -0.003 | -0.004 | -0.001 | -0.002 | -0.005 | -0.007 | -0.01 |
| GS | 0.02 | 0.022 | 0.021 | 0.02 | 0.018 | 0.019 | 0.024 | 0.024 | 0.027* | 0.029** | 0.03 |
| IBM | 0.024** | 0.02 | 0.02 | 0.014 | 0.012 | 0.015 | 0.008 | 0.012 | 0.007 | -0.001 | 0.02 |
| JPM | 0.015 | 0.015 | 0.016 | 0.013 | 0.01 | 0.014 | 0.009 | 0.012 | 0.01 | 0.011 | 0.00 |
| MSFT | 0.005 | 0.01 | 0.012 | 0.004 | -0.001 | -0.005 | -0.007 | -0.009 | -0.007 | -0.012 | -0.01 |
| NFLX | -0.2*** | -0.179*** | -0.136** | -0.107 | -0.125* | -0.107 | -0.158** | -0.135** | -0.131* | -0.145** | 0.09 |
| PG | 0.01 | 0.01 | 0.009 | 0.01 | 0.01 | 0.009 | 0.01* | 0.01 | 0.011* | 0.009 | 0.03 |
| SBUX | 0.012 | 0.007 | 0.006 | 0.007 | 0.006 | 0 | 0.003 | 0.002 | -0.001 | -0.006 | -0.01 |
| TSLA | -0.014 | -0.015 | -0.039 | -0.064 | -0.075 | -0.065 | -0.073 | -0.068 | -0.048 | -0.078 | 0.00 |
| TWTR | -0.167 | -0.186* | -0.182* | -0.208* | -0.165 | -0.163 | -0.193* | -0.276** | -0.23** | -0.24** | 0.05 |
| XOM | -0.028* | -0.005 | -0.004 | -0.004 | -0.004 | -0.003 | -0.004 | -0.004 | -0.004 | -0.004 | 0.00 |

Table 1.26: Intraday realized volatility vs Herding Rho (Non-directional W matrix, Log-odds)

| | Lags | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Firm | 1 day | 2 days | 3 days | 4 days | 5 days | 6 days | 7 days | 8 days | 9 days | 10 days | Rsq |
| AAPL | -0.0001 | 0.00019 | -0.00018 | -0.00017 | 0.00007 | -0.00014 | 0.00023 | -0.00019 | 0.00007 | 0.00008 | -0.03 |
| AMZN | -0.00023 | -0.00011 | 0.00047** | 0.00007 | 0.00003 | -0.00026 | -0.0001 | -0.00016 | -0.00016 | -0.00008 | 0.00 |
| BAC | -0.00013 | 0.00017 | -0.00007 | -0.00014 | -0.00017 | 0.00025 | 0.00011 | 0.00016 | -0.00004 | -0.00007 | -0.03 |
| BIDU | -0.00009 | -0.00012 | -0.00006 | -0.00011 | -0.00001 | 0.0001 | -0.00013 | -0.00009 | -0.00019* | -0.00017 | 0.01 |
| C | 0.00012 | 0.00008 | -0.00006 | -0.00001 | -0.00011 | -0.00015 | -0.0001 | 0.00036*** | 0.00003 | 0.00016 | 0.05 |
| CMG | 0.00004 | -0.00004 | 0.00002 | 0 | 0.00003 | -0.00003 | 0.00003 | -0.00007 | -0.00006 | 0.00012 | -0.02 |
| EBAY | -0.00002 | -0.00003 | -0.00002 | -0.00001 | -0.00005 | -0.00002 | -6e-05* | -6e-05** | -0.00005 | 0 | 0.01 |
| F | 0.00086 | 0.00471*** | 0.00137 | -0.00171 | -0.00171 | 0.00386** | 0.00441** | -0.0002 | -0.00045 | 0.00138 | 0.04 |
| FB | -0.00065* | -0.00071** | -0.00032 | 0.00012 | -0.00002 | -0.00021 | -0.00069* | 0.00006 | -0.00024 | -0.00028 | 0.03 |
| GE | 0.00002 | 0.00003 | 0.00011** | -0.00001 | 8e-05* | -0.00003 | 0 | -0.00005 | -0.00006 | -0.00007 | 0.02 |
| GILD | -0.00005 | 0.00023 | -0.00004 | -0.00029 | 0.00001 | 0.00008 | 0.00004 | -0.00033 | -0.00015 | 0.00003 | -0.03 |
| GM | -0.00009 | -0.00006 | -0.00016* | -0.00016* | 0.00007 | -0.00004 | -0.00004 | -0.00014 | -0.00015 | 0 | 0.01 |
| GS | 0.00004 | 0 | 0.00003 | 0.00003 | 0 | 0.00005 | 8e-05** | 0.00006 | 0.00006 | 0.00006 | 0.04 |
| IBM | -5.539 | -3.82306 | -3.92373 | -5.76143 | -4.45101 | -5.20455 | 2.72137 | -4.5295 | -4.86907 | 7.03323 | -0.03 |
| JPM | 5e-05* | 5e-05* | 0.00003 | -0.00001 | -0.00001 | -0.00001 | 0 | -0.00002 | 0.00001 | 0.00002 | -0.01 |
| MSFT | 0.00001 | 0.00001 | -0.00001 | 0 | 0 | 0.00001 | 0.00003 | 0.00002 | -0.00001 | 0.00001 | -0.03 |
| NFLX | -0.00074** | -0.00043 | -0.00034 | 0.00011 | 0.00019 | 0.00015 | -0.00054* | -0.00038 | -0.00002 | 0.00019 | 0.02 |
| PG | 0.00001 | 0.00001 | 0 | 0 | 0 | -0.00001 | -0.00002 | 0 | 5e-05*** | -0.00001 | 0.00 |
| SBUX | 0 | 0.00002 | 0 | -0.00002 | 0 | -0.00002 | -0.00002 | -0.00003 | -6e-05** | -6e-05** | 0.04 |
| TSLA | -0.00245* | -0.00173 | -0.00045 | -0.00136 | -0.00104 | -0.0007 | -0.00003 | -0.00031 | 0.00026 | -0.00123 | -0.01 |
| TWTR | 0 | -0.00021 | 0.00004 | -0.00033 | -0.00022 | -0.00052 | -0.00029 | -0.00004 | 0.00063* | 0.0003 | -0.01 |
| XOM | -0.00001 | -0.00001 | -0.00002 | -0.00002 | -0.00001 | 0 | 0 | 0 | 0.00002 | 0.00002 | -0.02 |

Table 1.27: Intraday realized volatility vs Herding Rho (Directional W matrix, Log-odds)

| | Lags | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Firm | 1 day | 2 days | 3 days | 4 days | 5 days | 6 days | 7 days | 8 days | 9 days | 10 days | Rsq |
| AAPL | -0.00015 | 0.00013 | -0.00018 | -0.00006 | 0.00014 | -0.00012 | 0.00041** | -0.00015 | 0.00016 | 0.0002 | 0.01 |
| AMZN | -0.00014 | 0.00004 | -0.00022 | 0.00006 | -0.00019 | -0.00025 | -0.00014 | -0.00024 | -0.00026 | -0.00001 | 0.01 |
| BAC | 0.00014 | 0.00039*** | 0.00001 | -0.00007 | -0.00007 | 0.00006 | 0.00005 | -0.00006 | -0.00007 | 0.00001 | 0.02 |
| BIDU | 0.00031** | 0.00003 | 0.00006 | 0.00006 | 0.00021* | 0.00022* | 0.00009 | -0.00006 | 0.00005 | 0.00011 | 0.03 |
| C | 0.00121*** | -0.00004 | 0.00012 | 0.00019 | 0.0002 | -0.00029 | 0.00009 | -0.00005 | 0.0003 | 0.00205*** | 0.24 |
| CMG | 0.00018 | 0.00012 | 0.00021 | 0.00016 | 0.00053*** | 0.00018 | 0.00025* | 0.00013 | 0.00044*** | 0.00017 | 0.07 |
| EBAY | 0.00023* | -0.00002 | -0.00002 | -0.00001 | 0.00018 | 0.00012 | 0.00023* | 0.00006 | 0.00003 | 0.00011 | 0.00 |
| F | -0.00103 | -0.00051 | -0.00108 | -0.00052 | -0.00029 | -0.00076 | 0.01142*** | -0.00062 | -0.00027 | -0.00035 | 0.13 |
| FB | -0.00037 | -0.00041* | -0.00012 | -0.0001 | -0.00034 | -0.00012 | -0.00049** | -0.00013 | -0.00029 | -0.00024 | 0.06 |
| GE | -0.00002 | -0.00001 | -0.00001 | -0.00002 | 0.00006 | 0.00002 | 0.00002 | 0.00004 | 8e-05* | 0.00002 | 0.00 |
| GILD | 0.00008 | 0.00012 | 0 | -0.00031 | 0.00008 | 0.00011 | 0.00011 | -0.00017 | 0.00012 | -0.00029 | -0.03 |
| GM | -0.00013 | -0.00007 | -0.00014 | -0.00028* | 0.00047*** | -0.00017 | -0.00006 | -0.00014 | -0.00015 | 0.00038** | 0.04 |
| GS | 0.00003 | 0.00002 | 0.00008 | 0 | -0.00001 | -0.00003 | 0.00006 | -0.00003 | 0.00002 | -0.00007 | -0.03 |
| IBM | -5.01 | -4.39732 | -4.94575 | -4.86817 | -3.22349 | -3.24844 | -4.83103 | -4.9468 | -4.25885 | -5.06123 | -0.04 |
| JPM | -0.00004 | 0.00004 | -0.00007 | -0.00008 | -0.00008 | -0.0001 | -0.0001 | -0.00003 | -0.00001 | -0.00008 | -0.03 |
| MSFT | 0.00005 | 0.00001 | 0.00005 | -0.00001 | 0 | 0 | 0.00005 | 0.00006 | 0.00002 | 0.00003 | -0.02 |
| NFLX | -0.0002 | 0.00003 | -0.00002 | 0.00005 | 0.00014 | 0.00004 | -0.00036 | -0.00008 | -0.00006 | 0.00016 | -0.03 |
| PG | -0.00001 | -0.00003 | -0.00002 | 0.00001 | -0.00001 | -0.00001 | 0 | -0.00002 | -0.00001 | -0.00002 | -0.03 |
| SBUX | 0.00002 | -0.00001 | 0.00003 | -0.00004 | 0.00001 | -0.00005 | 0.00003 | -0.00007 | -0.00003 | -0.00007 | -0.02 |
| TSLA | -0.00099 | 0.00014 | -0.00009 | 0.00009 | -0.00041 | -0.00075 | 0.00076 | 0.00003 | -0.00021 | -0.00117 | -0.03 |
| TWTR | 0.00028 | -0.00008 | 0.00066*** | 0.00012 | -0.00008 | -0.00023 | -6e-04** | -0.00016 | -0.00018 | -0.00022 | 0.03 |
| XOM | -0.00002 | -0.00001 | 0.00005 | 0.00002 | -0.00002 | -0.00001 | 0 | 0.00001 | -0.00001 | 0.00001 | -0.04 |

Table 1.28: Intraday realized volatility vs Herding Rho (Directional W matrix, Probability)

| | Lags | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Firm | 1 day | 2 days | 3 days | 4 days | 5 days | 6 days | 7 days | 8 days | 9 days | 10 days | Rsq |
| AAPL | -0.00041* | -0.00008 | -0.00011 | -0.00028 | 0 | -0.00003 | 0.00011 | -0.00013 | -0.00001 | 0.00021 | -0.01 |
| AMZN | -0.00003 | -0.00008 | -0.00012 | 0.00015 | 0.00003 | -0.00008 | -0.00033* | -0.00018 | 0.00007 | 0.00008 | -0.02 |
| BAC | 0.00013 | 0.00025* | 0.00025* | -0.00019 | -0.00015 | 0.00001 | 0.00001 | -0.0001 | 0.0001 | -0.00004 | 0.00 |
| BIDU | 0.00032** | 0.00007 | 0.0001 | -0.00001 | 0.00014 | 0.00014 | 0.00006 | -0.00012 | 0.00004 | 0.00016 | 0.01 |
| C | 0.00085*** | -0.00003 | -0.00002 | 0.00016 | 0.00033 | -0.00081*** | 0.00027 | -0.00013 | 0.00058** | 0.00271*** | 0.33 |
| CMG | 0.00012 | 0.00007 | 0.00015 | 0.0001 | 0.00034** | 0.0002 | 0.00015 | 0.00015 | 0.00012 | 0.00012 | 0.01 |
| EBAY | 0.00033*** | -0.00007 | -0.00001 | -0.00002 | -0.00002 | -0.00002 | -0.00002 | -0.00004 | -0.0001 | -0.00003 | -0.01 |
| F | -0.00049 | -0.0006 | -0.00084 | -0.00067 | -0.00065 | -0.00075 | 0.00336* | -0.00076 | -0.00065 | -0.00061 | -0.02 |
| FB | -0.00068** | -0.0004 | -0.00056* | 0.00015 | -0.0002 | -0.00009 | -0.00001 | -0.00041 | -0.00006 | -0.00002 | 0.02 |
| GE | -0.00002 | 0.00003 | 0 | 0 | 9e-05** | 7e-05* | 9e-05** | 7e-05* | 0.00012*** | 7e-05* | 0.10 |
| GILD | -0.00016 | 0.00024 | 0.00013 | -0.00038 | 0.00011 | 0.00046 | 0.00031 | 0.00004 | -0.00001 | -0.00014 | -0.03 |
| GM | -0.00014 | -0.00016 | -0.00005 | -0.00021 | 0.00053*** | -0.00012 | -0.00012 | -0.0002 | -0.00013 | 0.00042** | 0.04 |
| GS | 0.00003 | 0.00002 | 0.00011 | 0.00009 | 0.00001 | -0.00005 | 0.00006 | -0.00004 | 0.00002 | -0.00008 | -0.03 |
| IBM | -3.82237 | -3.77967 | -4.0146 | -3.92736 | -2.68711 | -2.67443 | -3.62707 | -3.74001 | -3.40904 | -3.46427 | -0.04 |
| JPM | 0.00017 | 0.00042** | 0.00036** | 0 | 0.00008 | 0.0001 | 0.00012 | 0.00019 | 0.00009 | -0.00003 | 0.02 |
| MSFT | 0.00008 | -0.00003 | 0.00003 | 0.00007 | 0.00002 | -0.00001 | -0.00002 | 0.00006 | -0.00004 | 0.00003 | -0.01 |
| NFLX | -0.00012 | -0.00002 | -0.00002 | -0.00002 | 0.00015 | 0.00015 | -0.00022 | 0.00029 | -0.00026 | -0.00017 | -0.02 |
| PG | -0.00001 | -0.00002 | -0.00001 | -0.00001 | -0.00001 | -0.00001 | -0.00002 | -0.00002 | 0 | -0.00003 | -0.02 |
| SBUX | -0.00001 | 0 | 0 | -0.00002 | -0.00002 | -0.00001 | -0.00001 | -0.00002 | 0.00001 | -0.00001 | -0.03 |
| TSLA | -0.00071 | -0.00004 | -0.0007 | 0.00091 | -0.00019 | 0.00199* | 0.0016 | 0.00128 | 0.00095 | 0.00076 | -0.01 |
| TWTR | 0.00045 | 0.00026 | 0.00057* | -0.00004 | 0.00009 | -0.00028 | 0.00016 | 0.00028 | 0.00009 | 0.00038 | 0.00 |
| XOM | 0 | 0 | 0.00002 | 0 | 0 | 0 | 0 | 0 | -0.00001 | 0.00001 | -0.04 |

Table 1.29: Realized volatility (30 days) vs Herding Rho (Non-directional W matrix, Probability)

| | Lags | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Firm | 1 day | 2 days | 3 days | 4 days | 5 days | 6 days | 7 days | 8 days | 9 days | 10 days | Rsq |
| AAPL | -0.00047 | -0.00026 | -0.00005 | -0.00018 | -0.00008 | -0.00023 | 0.00002 | -0.00004 | -0.00021 | 0.00037 | -0.01 |
| AMZN | -0.00018 | -0.00015 | 0.00042* | 0.00014 | 0.00003 | 0.00012 | -0.00027 | 0 | 0.00024 | -0.00002 | -0.01 |
| BAC | 0.00002 | 0.00025 | 0.0002 | -0.00035* | -0.00033* | 0.00023 | -0.00002 | -0.00012 | -0.0001 | -0.00022 | 0.01 |
| BIDU | 0.00004 | 0.00001 | -0.00002 | -0.00007 | 0.00006 | 0.00009 | 0.00008 | -0.00007 | -0.00006 | -0.0001 | -0.02 |
| C | 0.00006 | 0.00004 | -0.00009 | -0.00006 | -0.00007 | -0.00011 | -0.00007 | 0.00011 | 0.00009 | 0.00029*** | 0.02 |
| CMG | -0.00002 | -0.00006 | 0.00001 | -0.00005 | 0.00006 | 0.00002 | 0.00008 | -0.00007 | -0.00002 | 0.00025*** | 0.02 |
| EBAY | -0.00003 | -0.00004 | -0.00001 | -0.00002 | -0.00001 | -0.00003 | -6e-05** | -0.00004 | -6e-05* | 0.00001 | 0.01 |
| F | 0.00144 | 0.00342** | -0.0001 | -0.00101 | -0.00175 | 0.00224 | -0.00035 | -0.00033 | -0.00125 | -0.00142 | 0.00 |
| FB | -0.00115*** | -0.00068 | -0.00061 | -0.00025 | -0.00039 | -0.00028 | -0.00005 | -0.00034 | -0.00062 | -0.00001 | 0.05 |
| GE | 0.00001 | 0.00004 | 0 | -0.00001 | 0.00004 | 0.00001 | 0.00002 | -0.00003 | -0.00001 | -0.00005 | -0.03 |
| GILD | -0.00013 | -0.00001 | -0.00009 | -0.00025 | 0.00003 | 0.0001 | 0.00016 | -0.00011 | -0.00004 | -0.00011 | -0.04 |
| GM | -0.00009 | 0.00002 | -0.00005 | -0.00015* | 0.00003 | -0.00007 | -0.00005 | -0.00013 | -0.00015* | -0.00003 | 0.01 |
| GS | 0.00003 | 0.00003 | 7e-05* | 0.00006 | 0.00006 | 0.00006 | 9e-05** | 0.00005 | 0.00011*** | 7e-05* | 0.11 |
| IBM | -6.07103 | -3.76977 | -3.9592 | -4.40501 | -4.29187 | -6.26722 | 14.33849 | -6.13744 | -4.22169 | -4.35017 | -0.03 |
| JPM | 0.00002 | 0.00002 | 0 | -0.00003 | -0.00003 | -0.00001 | 0 | -0.00001 | 0 | 0 | -0.02 |
| MSFT | 0 | 0.00004 | 0.00002 | 0.00002 | 0.00003 | 0.00004 | 0.00002 | 0.00003 | 0.00003 | -0.00001 | -0.01 |
| NFLX | -0.0005 | -0.00028 | -0.00006 | -0.00005 | 0.00002 | 0.00049 | 0.00015 | 0.0001 | -0.00034 | -0.00021 | -0.01 |
| PG | 0.00001 | 0.00001 | 0 | 0.00001 | 0.00001 | 0 | -0.00001 | 0 | 0.00001 | -0.00001 | -0.01 |
| SBUX | 0.00001 | 0 | -0.00001 | -0.00003 | 0 | -0.00001 | -0.00001 | -4e-05* | -4e-05* | -4e-05** | 0.04 |
| TSLA | -0.00225 | 0.00038 | -0.00052 | 0.00018 | -0.00181 | 0.00204 | 0.00107 | 0 | 0.00172 | -0.00126 | -0.01 |
| TWTR | -0.00031 | 0.00029 | -0.00009 | -0.00059 | -0.00009 | -0.0005 | -0.00029 | -0.00005 | 0.00022 | 0.00012 | -0.02 |
| XOM | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.00001 | 0 | -0.03 |

Table 1.30: R-squared and partial R-squared in the regression on lagged Volatility and Herding Rho (directional W matrix)

| | Intraday Vol | | 10-day Vol | | 14-day Vol | | 30-day Vol | |
|---|---|---|---|---|---|---|---|---|
| Firms | Adjusted $R^2$ | Partial $R^2$ | Adjusted $R^2$ | Partial $R^2$ | Adjusted $R^2$ | Partial $R^2$ | Adjusted $R^2$ | Partial $R^2$ |
| Apple Inc. | 0.0396 | 0.0556 | 0.8007 | 0.0231 | 0.8698 | 0.0290 | 0.9402 | 0.0177 |
| Amazon.com, Inc. | 0.1740 | 0.0385 | 0.7625 | 0.0078 | 0.8546 | 0.0123 | 0.9228 | 0.0124 |
| Bank of America Corporation | 0.0223 | 0.0583 | 0.8473 | 0.0192 | 0.9078 | 0.0090 | 0.9623 | 0.0084 |
| Baidu, Inc. | 0.3124 | 0.0530 | 0.7720 | 0.0499 | 0.8471 | 0.0423 | 0.9429 | 0.0429 |
| Citigroup Inc. | 0.2325 | 0.2652 | 0.8496 | 0.0361 | 0.9086 | 0.0284 | 0.9652 | 0.0353 |
| Chipotle Mexican Grill, Inc. | 0.1395 | 0.0584 | 0.7534 | 0.0088 | 0.8241 | 0.0056 | 0.9189 | 0.0084 |
| eBay Inc. | 0.0783 | 0.0394 | 0.7427 | 0.0145 | 0.8196 | 0.0094 | 0.9123 | 0.0073 |
| Ford Motor Company | 0.1316 | 0.1678 | 0.7661 | 0.0110 | 0.8204 | 0.0129 | 0.9272 | 0.0221 |
| Facebook, Inc. | 0.2292 | 0.0573 | 0.8302 | 0.0206 | 0.8869 | 0.0176 | 0.9566 | 0.0171 |
| General Electric Company | 0.1194 | 0.0338 | 0.8337 | 0.0135 | 0.8839 | 0.0112 | 0.9516 | 0.0085 |
| Gilead Sciences, Inc. | 0.0230 | 0.0160 | 0.7690 | 0.0070 | 0.8203 | 0.0108 | 0.9151 | 0.0132 |
| General Motors Company | 0.0826 | 0.0925 | 0.7534 | 0.0188 | 0.8185 | 0.0060 | 0.9261 | 0.0189 |
| Goldman Sachs Group, Inc. | 0.2322 | 0.0613 | 0.8389 | 0.0110 | 0.9039 | 0.0098 | 0.9672 | 0.0206 |
| International Business Machines | -0.0423 | 0.0013 | 0.8027 | 0.0145 | 0.8610 | 0.0110 | 0.9349 | 0.0195 |
| J P Morgan Chase & Co | 0.4477 | 0.0512 | 0.8261 | 0.0256 | 0.8995 | 0.0227 | 0.9596 | 0.0179 |
| Microsoft Corporation | 0.2948 | 0.0277 | 0.7809 | 0.0154 | 0.8516 | 0.0268 | 0.9305 | 0.0131 |
| Netflix, Inc. | 0.3376 | 0.0510 | 0.7518 | 0.0173 | 0.8124 | 0.0164 | 0.9033 | 0.0248 |
| Procter & Gamble Company | 0.1057 | 0.0182 | 0.8004 | 0.0081 | 0.8682 | 0.0134 | 0.9508 | 0.0246 |
| Starbucks Corporation | 0.2341 | 0.0656 | 0.8136 | 0.0280 | 0.8808 | 0.0217 | 0.9562 | 0.0388 |
| Tesla, Inc. | -0.0038 | 0.0363 | 0.7798 | 0.0985 | 0.8382 | 0.1186 | 0.9369 | 0.0897 |
| Twitter, Inc. | 0.2758 | 0.1556 | 0.7673 | 0.0320 | 0.8162 | 0.0150 | 0.9083 | 0.0106 |
| Exxon Mobil Corporation | 0.6131 | 0.0291 | 0.8658 | 0.0162 | 0.9075 | 0.0205 | 0.9730 | 0.0205 |

Table 1.31: R-squared and partial R-squared in the regression on lagged Volatility and Herding Rho (non-directional W matrix)

| | Intraday Vol | | 10-day Vol | | 14-day Vol | | 30-day Vol | |
|---|---|---|---|---|---|---|---|---|
| Firms | Adjusted $R^2$ | Partial $R^2$ | Adjusted $R^2$ | Partial $R^2$ | Adjusted $R^2$ | Partial $R^2$ | Adjusted $R^2$ | Partial $R^2$ |
| Apple Inc. | 0.0009 | 0.0175 | 0.7988 | 0.0139 | 0.8698 | 0.0287 | 0.9402 | 0.0178 |
| Amazon.com, Inc. | 0.1755 | 0.0403 | 0.7637 | 0.0131 | 0.8552 | 0.0162 | 0.9228 | 0.0128 |
| Bank of America Corporation | -0.0188 | 0.0187 | 0.8468 | 0.0159 | 0.9084 | 0.0148 | 0.9624 | 0.0111 |
| Baidu, Inc. | 0.3044 | 0.0420 | 0.7637 | 0.0155 | 0.8426 | 0.0140 | 0.9414 | 0.0191 |
| Citigroup Inc. | 0.0501 | 0.0906 | 0.8483 | 0.0280 | 0.9090 | 0.0325 | 0.9652 | 0.0367 |
| Chipotle Mexican Grill, Inc. | 0.1109 | 0.0270 | 0.7542 | 0.0120 | 0.8267 | 0.0203 | 0.9187 | 0.0059 |
| eBay Inc. | 0.0884 | 0.0499 | 0.7425 | 0.0139 | 0.8202 | 0.0129 | 0.9123 | 0.0075 |
| Ford Motor Company | 0.0426 | 0.0826 | 0.7654 | 0.0082 | 0.8205 | 0.0134 | 0.9263 | 0.0098 |
| Facebook, Inc. | 0.2211 | 0.0474 | 0.8304 | 0.0220 | 0.8860 | 0.0095 | 0.9563 | 0.0117 |
| General Electric Company | 0.1533 | 0.0709 | 0.8345 | 0.0178 | 0.8856 | 0.0260 | 0.9518 | 0.0131 |
| Gilead Sciences, Inc. | 0.0199 | 0.0128 | 0.7685 | 0.0050 | 0.8205 | 0.0119 | 0.9149 | 0.0109 |
| General Motors Company | 0.0342 | 0.0446 | 0.7550 | 0.0251 | 0.8191 | 0.0092 | 0.9258 | 0.0150 |
| Goldman Sachs Group, Inc. | 0.2510 | 0.0843 | 0.8388 | 0.0106 | 0.9044 | 0.0151 | 0.9673 | 0.0232 |
| International Business Machines | -0.0319 | 0.0112 | 0.8031 | 0.0169 | 0.8613 | 0.0134 | 0.9350 | 0.0211 |
| J P Morgan Chase & Co | 0.4480 | 0.0518 | 0.8274 | 0.0326 | 0.8994 | 0.0222 | 0.9595 | 0.0164 |
| Microsoft Corporation | 0.2915 | 0.0230 | 0.7805 | 0.0134 | 0.8502 | 0.0177 | 0.9311 | 0.0214 |
| Netflix, Inc. | 0.3451 | 0.0617 | 0.7547 | 0.0289 | 0.8166 | 0.0381 | 0.9036 | 0.0277 |
| Procter & Gamble Company | 0.1096 | 0.0224 | 0.8015 | 0.0137 | 0.8699 | 0.0265 | 0.9508 | 0.0247 |
| Starbucks Corporation | 0.2362 | 0.0681 | 0.8132 | 0.0261 | 0.8820 | 0.0315 | 0.9562 | 0.0404 |
| Tesla, Inc. | 0.0070 | 0.0467 | 0.7825 | 0.1094 | 0.8413 | 0.1356 | 0.9368 | 0.0888 |
| Twitter, Inc. | 0.2234 | 0.0944 | 0.7661 | 0.0271 | 0.8177 | 0.0229 | 0.9091 | 0.0186 |
| Exxon Mobil Corporation | 0.6073 | 0.0147 | 0.8655 | 0.0144 | 0.9064 | 0.0092 | 0.9732 | 0.0288 |

Table 1.32: R-squared and partial R-squared in the regression on lagged Volatility and Herding Rho

Directional W matrix, errors from regression on multiple News Polarity lags

| | Intraday Vol | | 10-day Vol | | 14-day Vol | | 30-day Vol | |
|---|---|---|---|---|---|---|---|---|
| Firms | Adjusted R | Partial R | Adjusted R | Partial R | Adjusted R | Partial R | Adjusted R | Partial R |
| Apple Inc. | 0.0065 | 0.0278 | 0.8003 | 0.0217 | 0.8703 | 0.0199 | 0.9404 | 0.0224 |
| Amazon.com, Inc. | 0.1408 | 0.0794 | 0.7636 | 0.0497 | 0.8584 | 0.0626 | 0.9262 | 0.0715 |
| Bank of America Corporation | -0.0169 | 0.0256 | 0.8464 | 0.0139 | 0.9084 | 0.0166 | 0.9624 | 0.0111 |
| Facebook, Inc. | 0.2164 | 0.0474 | 0.8310 | 0.0255 | 0.8862 | 0.0152 | 0.9562 | 0.0140 |
| General Electric Company | 0.1317 | 0.1559 | 0.8333 | 0.0140 | 0.8844 | 0.0168 | 0.9520 | 0.0153 |
| Gilead Sciences, Inc. | 0.0710 | 0.0708 | 0.7703 | 0.0153 | 0.8213 | 0.0182 | 0.9166 | 0.0311 |
| Netflix, Inc. | 0.4018 | 0.1485 | 0.7651 | 0.0834 | 0.8231 | 0.0823 | 0.9100 | 0.0986 |
| Tesla, Inc. | -0.0092 | 0.0363 | 0.7723 | 0.1182 | 0.8289 | 0.1261 | 0.9353 | 0.0939 |
| Twitter, Inc. | 0.2389 | 0.0835 | 0.7686 | 0.0300 | 0.8174 | 0.0147 | 0.9079 | 0.0277 |
| Exxon Mobil Corporation | 0.6199 | 0.0526 | 0.8682 | 0.0350 | 0.9071 | 0.0183 | 0.9734 | 0.0381 |

Non-directional W matrix , errors from regression on multiple News Polarity lags)

| | Intraday Vol | | 10-day Vol | | 14-day Vol | | 30-day Vol | |
|---|---|---|---|---|---|---|---|---|
| Firms | Adjusted R | Partial R | Adjusted R | Partial R | Adjusted R | Partial R | Adjusted R | Partial R |
| Apple Inc. | -0.0095 | 0.0121 | 0.8006 | 0.0235 | 0.8695 | 0.0143 | 0.9399 | 0.0148 |
| Amazon.com, Inc. | 0.1533 | 0.0928 | 0.7648 | 0.0546 | 0.8585 | 0.0630 | 0.9264 | 0.0740 |
| Bank of America Corporation | 0.0146 | 0.0558 | 0.8457 | 0.0094 | 0.9085 | 0.0178 | 0.9630 | 0.0275 |
| Facebook, Inc. | 0.2284 | 0.0619 | 0.8311 | 0.0264 | 0.8865 | 0.0179 | 0.9562 | 0.0134 |
| General Electric Company | 0.1850 | 0.2077 | 0.8347 | 0.0225 | 0.8853 | 0.0246 | 0.9520 | 0.0173 |
| Gilead Sciences, Inc. | 0.0889 | 0.0887 | 0.7726 | 0.0252 | 0.8227 | 0.0263 | 0.9152 | 0.0148 |
| Netflix, Inc. | 0.3688 | 0.1015 | 0.7604 | 0.0650 | 0.8201 | 0.0666 | 0.9077 | 0.0747 |
| Tesla, Inc. | 0.0165 | 0.0608 | 0.7765 | 0.1343 | 0.8317 | 0.1408 | 0.9363 | 0.1077 |
| Twitter, Inc. | 0.3039 | 0.1618 | 0.7669 | 0.0230 | 0.8169 | 0.0124 | 0.9068 | 0.0158 |
| Exxon Mobil Corporation | 0.6225 | 0.0593 | 0.8654 | 0.0141 | 0.9068 | 0.0151 | 0.9731 | 0.0267 |

Table 1.33: ADF test statistics. News Polarity and Rho

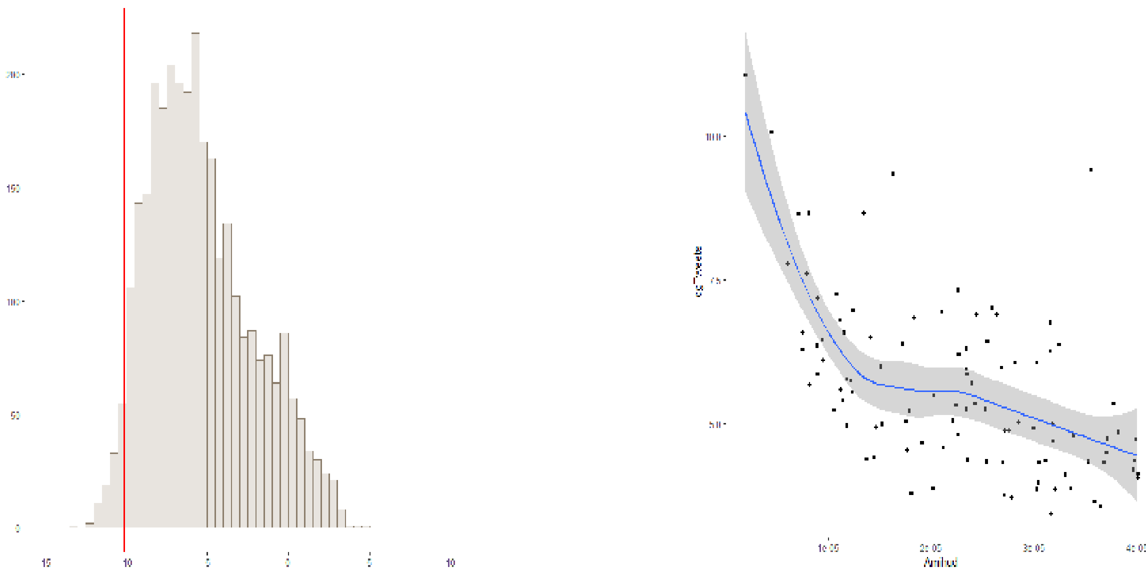
| Firm | News Polarity (n lags = 10) | Rho (n lags = 9) |
|---|---|---|
| Apple Inc. | -9.99 | -7.718 |
| Amazon.com, Inc. | -10.62 | -8.543 |
| Bank of America Corporation | -10.00 | -8.761 |
| Baidu, Inc. | -9.84 | -8.706 |
| Citigroup Inc. | -9.45 | -9.951 |
| Chipotle Mexican Grill, Inc. | -9.40 | -9.487 |
| eBay Inc. | -9.44 | -8.588 |
| Ford Motor Company | -9.64 | -9.120 |
| Facebook, Inc. | -9.94 | -7.185 |
| General Electric Company | -8.84 | -7.824 |
| Gilead Sciences, Inc. | -9.36 | -8.378 |
| General Motors Company | -10.10 | -9.608 |
| Goldman Sachs Group, Inc. | -9.18 | -8.674 |
| International Business Machines Corporation | -10.01 | -8.746 |
| J P Morgan Chase & Co | -9.79 | -9.244 |
| Microsoft Corporation | -9.50 | -8.424 |
| Netflix, Inc. | -10.64 | -8.903 |
| Procter & Gamble Company | -10.74 | -9.145 |
| Starbucks Corporation | -10.29 | -8.832 |
| Tesla, Inc. | -10.05 | -9.000 |
| Twitter, Inc. | -8.37 | -9.325 |
| Exxon Mobil Corporation | -9.23 | -8.847 |

Table 1.34: R-squared and partial R-squared in the regression of differenced volatility on lagged differenced volatility and Herding Rho (non-directional W matrix)

| | 10-day Vol | | 14-day Vol | | 30-day Vol | |
|---|---|---|---|---|---|---|
| Firms | Adjusted R | Partial R | Adjusted R | Partial R | Adjusted R | Partial R |
| Apple Inc. | -0.0066 | 0.0114 | -0.0026 | 0.0287 | 0.0047 | 0.0195 |
| Amazon.com, Inc. | 0.0087 | 0.0119 | 0.0024 | 0.0169 | 0.0027 | 0.0157 |
| Bank of America Corporation | 0.0061 | 0.0095 | 0.0202 | 0.0378 | 0.0078 | 0.0235 |
| Baidu, Inc. | -0.0058 | 0.0168 | -0.0073 | 0.0108 | -0.0017 | 0.0210 |
| Citigroup Inc. | 0.0059 | 0.0268 | 0.0207 | 0.0533 | 0.0070 | 0.0465 |
| Chipotle Mexican Grill, Inc. | -0.0037 | 0.0103 | 0.0046 | 0.0197 | -0.0007 | 0.0080 |
| eBay Inc. | -0.0058 | 0.0117 | -0.0020 | 0.0144 | -0.0096 | 0.0071 |
| Ford Motor Company | -0.0066 | 0.0110 | -0.0084 | 0.0133 | -0.0064 | 0.0078 |
| Facebook, Inc. | -0.0092 | 0.0218 | -0.0058 | 0.0127 | -0.0062 | 0.0114 |
| General Electric Company | 0.0199 | 0.0253 | 0.0270 | 0.0422 | 0.0074 | 0.0212 |
| Gilead Sciences, Inc. | -0.0098 | 0.0069 | -0.0023 | 0.0164 | -0.0055 | 0.0125 |
| General Motors Company | 0.0058 | 0.0244 | -0.0044 | 0.0077 | -0.0048 | 0.0139 |
| Goldman Sachs Group, Inc. | -0.0021 | 0.0093 | -0.0004 | 0.0162 | -0.0008 | 0.0177 |
| International Business Machines | 0.0007 | 0.0175 | 0.0046 | 0.0220 | 0.0129 | 0.0319 |
| J P Morgan Chase & Co | 0.0154 | 0.0185 | 0.0177 | 0.0408 | 0.0062 | 0.0298 |
| Microsoft Corporation | -0.0049 | 0.0157 | -0.0022 | 0.0199 | 0.0079 | 0.0265 |
| Netflix, Inc. | 0.0116 | 0.0318 | 0.0224 | 0.0408 | 0.0076 | 0.0255 |
| Procter & Gamble Company | 0.0041 | 0.0173 | 0.0156 | 0.0326 | 0.0103 | 0.0377 |
| Starbucks Corporation | 0.0002 | 0.0229 | 0.0121 | 0.0372 | 0.0100 | 0.0430 |
| Tesla, Inc. | 0.0094 | 0.1058 | 0.0153 | 0.1350 | -0.0103 | 0.0864 |
| Twitter, Inc. | 0.0011 | 0.0312 | 0.0038 | 0.0266 | 0.0040 | 0.0201 |
| Exxon Mobil Corporation | -0.0053 | 0.0142 | -0.0043 | 0.0130 | 0.0080 | 0.0280 |

Table 1.35: R-squared and partial R-squared in the regression of differenced volatility on lagged differenced volatility and Herding Rho (directional W matrix)

| | 10-day Vol | | 14-day Vol | | 30-day Vol | |
|---|---|---|---|---|---|---|
| Firms | Adjusted R | Partial R | Adjusted R | Partial R | Adjusted R | Partial R |
| Apple Inc. | 0.0057 | 0.0234 | 0.0001 | 0.0313 | 0.0063 | 0.0212 |
| Amazon.com, Inc. | 0.0001 | 0.0033 | -0.0027 | 0.0119 | -0.0008 | 0.0122 |
| Bank of America Corporation | 0.0084 | 0.0119 | 0.0134 | 0.0311 | 0.0030 | 0.0188 |
| Baidu, Inc. | 0.0161 | 0.0381 | 0.0094 | 0.0272 | 0.0158 | 0.0381 |
| Citigroup Inc. | 0.0141 | 0.0348 | 0.0182 | 0.0509 | 0.0068 | 0.0463 |
| Chipotle Mexican Grill, Inc. | -0.0070 | 0.0070 | -0.0080 | 0.0073 | 0.0014 | 0.0101 |
| eBay Inc. | -0.0038 | 0.0137 | -0.0053 | 0.0112 | -0.0083 | 0.0085 |
| Ford Motor Company | -0.0051 | 0.0125 | -0.0095 | 0.0122 | 0.0046 | 0.0187 |
| Facebook, Inc. | -0.0090 | 0.0220 | 0.0055 | 0.0237 | -0.0011 | 0.0164 |
| General Electric Company | 0.0028 | 0.0083 | 0.0083 | 0.0238 | 0.0000 | 0.0139 |
| Gilead Sciences, Inc. | -0.0072 | 0.0094 | -0.0042 | 0.0146 | -0.0025 | 0.0155 |
| General Motors Company | -0.0014 | 0.0173 | -0.0085 | 0.0036 | -0.0012 | 0.0175 |
| Goldman Sachs Group, Inc. | 0.0042 | 0.0156 | -0.0011 | 0.0155 | 0.0066 | 0.0249 |
| International Business Machines | -0.0014 | 0.0154 | 0.0056 | 0.0230 | 0.0007 | 0.0200 |
| J P Morgan Chase & Co | 0.0071 | 0.0102 | 0.0190 | 0.0420 | 0.0096 | 0.0330 |
| Microsoft Corporation | -0.0013 | 0.0192 | 0.0084 | 0.0302 | 0.0058 | 0.0245 |
| Netflix, Inc. | -0.0014 | 0.0190 | -0.0022 | 0.0167 | 0.0086 | 0.0264 |
| Procter & Gamble Company | -0.0072 | 0.0062 | 0.0012 | 0.0185 | 0.0005 | 0.0281 |
| Starbucks Corporation | 0.0030 | 0.0255 | 0.0009 | 0.0263 | 0.0082 | 0.0412 |
| Tesla, Inc. | -0.0031 | 0.0945 | -0.0084 | 0.1142 | -0.0091 | 0.0875 |
| Twitter, Inc. | 0.0077 | 0.0376 | -0.0040 | 0.0190 | -0.0011 | 0.0152 |
| Exxon Mobil Corporation | -0.0032 | 0.0162 | 0.0080 | 0.0250 | 0.0024 | 0.0225 |




(a) Distribuion of log of Amihud illiquidity

(b) Number of tweets vs Amihud measure in the left cut-off

Figure 1.1: Stocks' Amihud measure of illiquidity and social media coverage
Panel (a) presents the distribution of log of Amuhud measure calculated according to Amihud (2002) and averaged over the three years; red line represents cut-off for the most liquid 100 stocks traded on NYSE, NASDAQ, or AMEX.(Most liquid stocks have smallest Amihud number, and are located in the left tail of distribution). Panel (b) presents relationship between the log of number of tweets posted during January 2014 (first month in the time window considered in the study) for these 100 stocks and their Amihud illiquidity. measure.

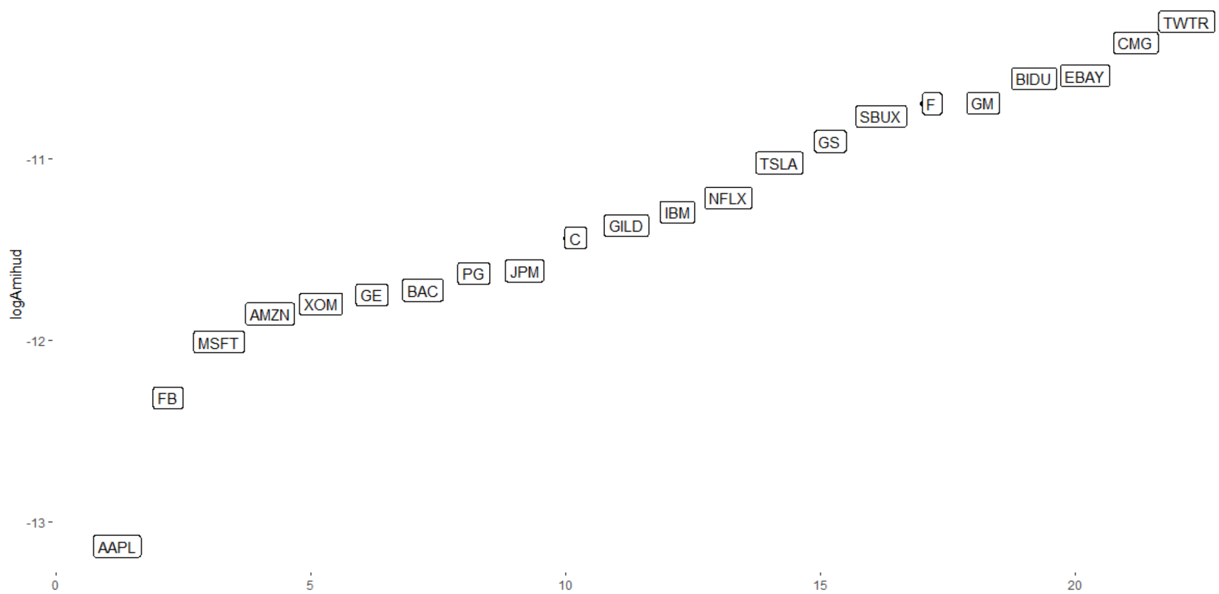



Figure 1.2: Stocks with sufficient continuous tweet data and their log-Amihud illiquidity
The 22 stocks are chosen with the log-Amihud measure evenly covering the left cut-off of the distribution as presented on Figure 1(a).

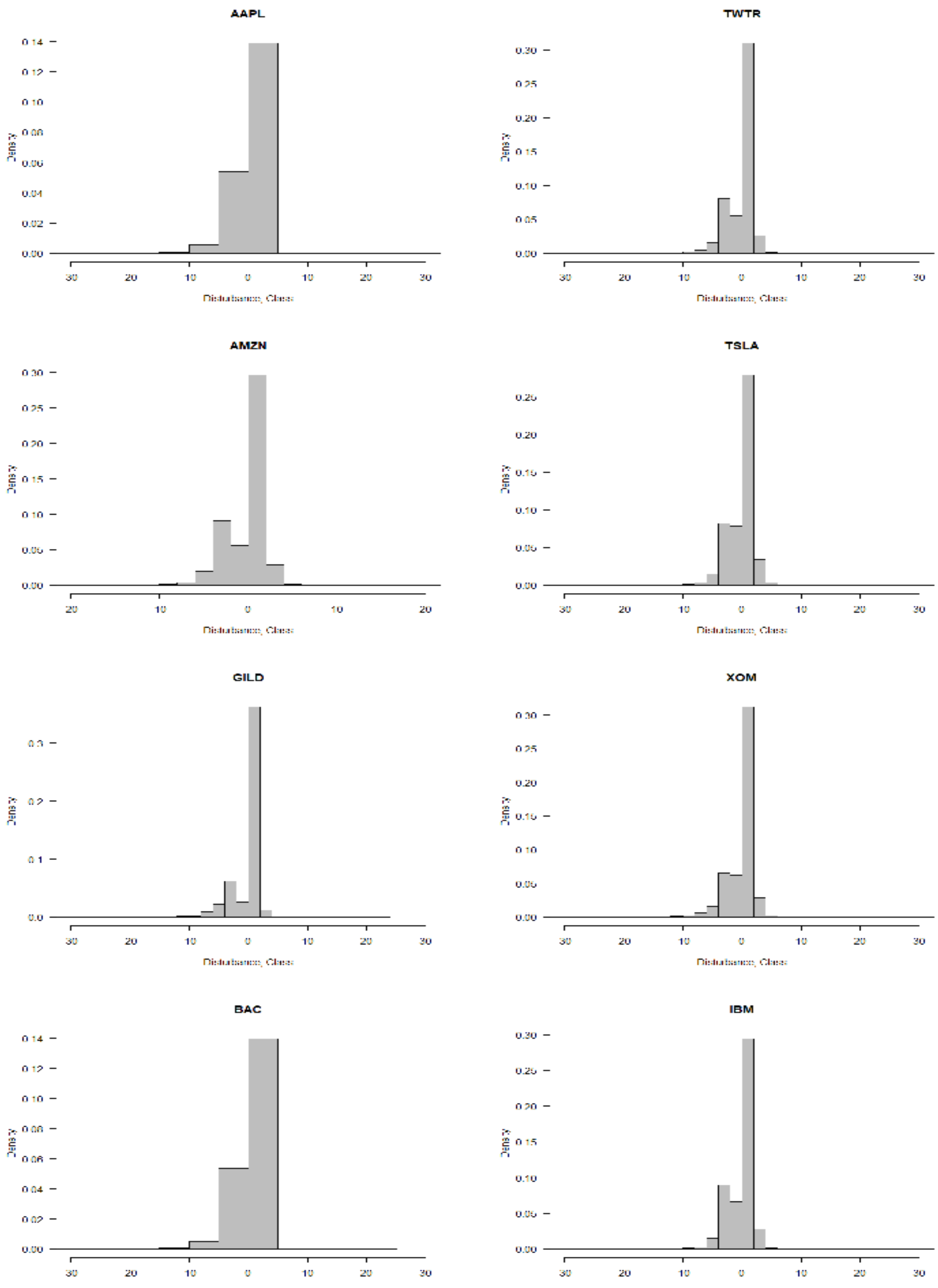



Figure 1.3: Histograms of disturbances as defined by Equation (8)

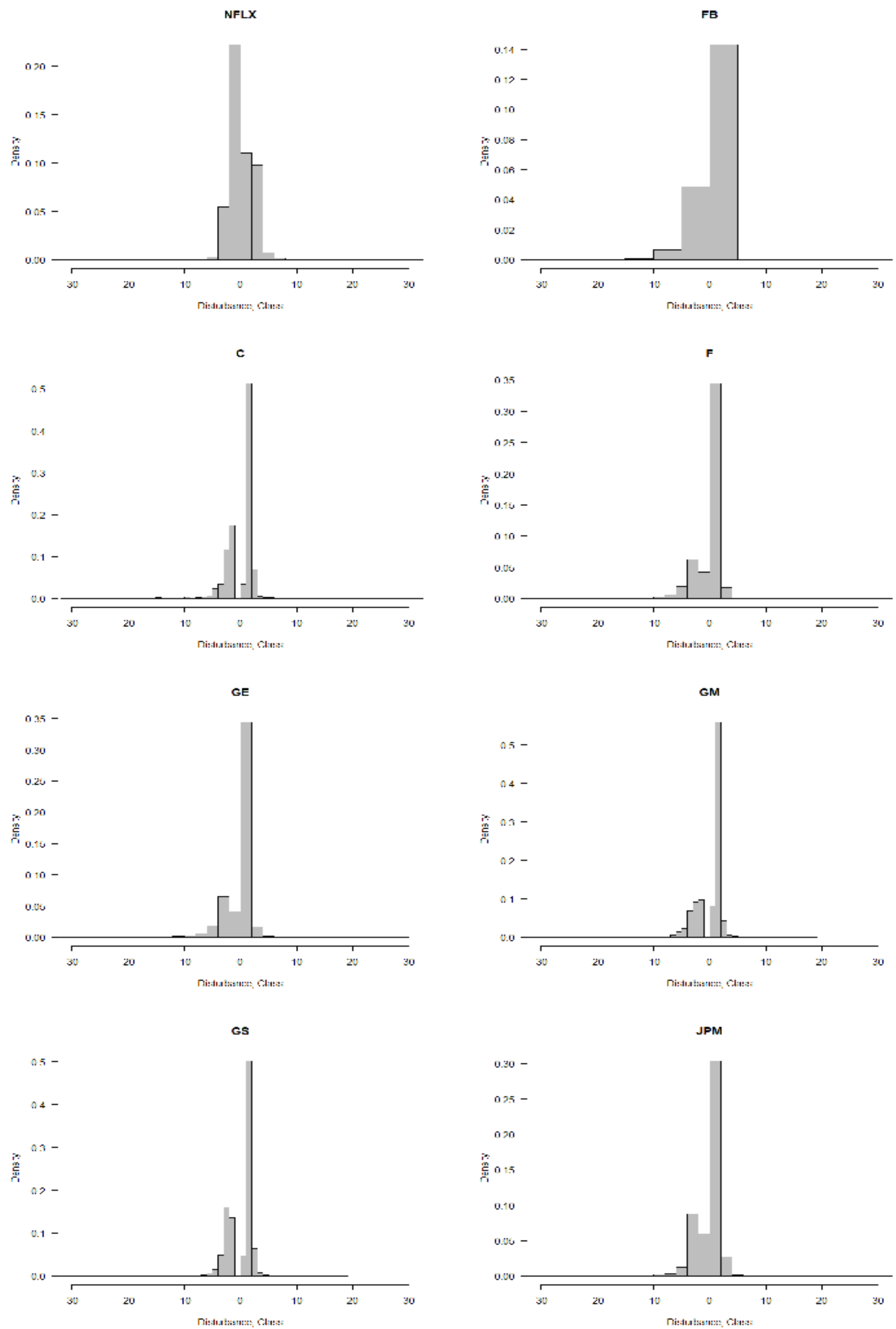



Figure 1.4: Histograms of disturbances as defined by Equation (8). Continued

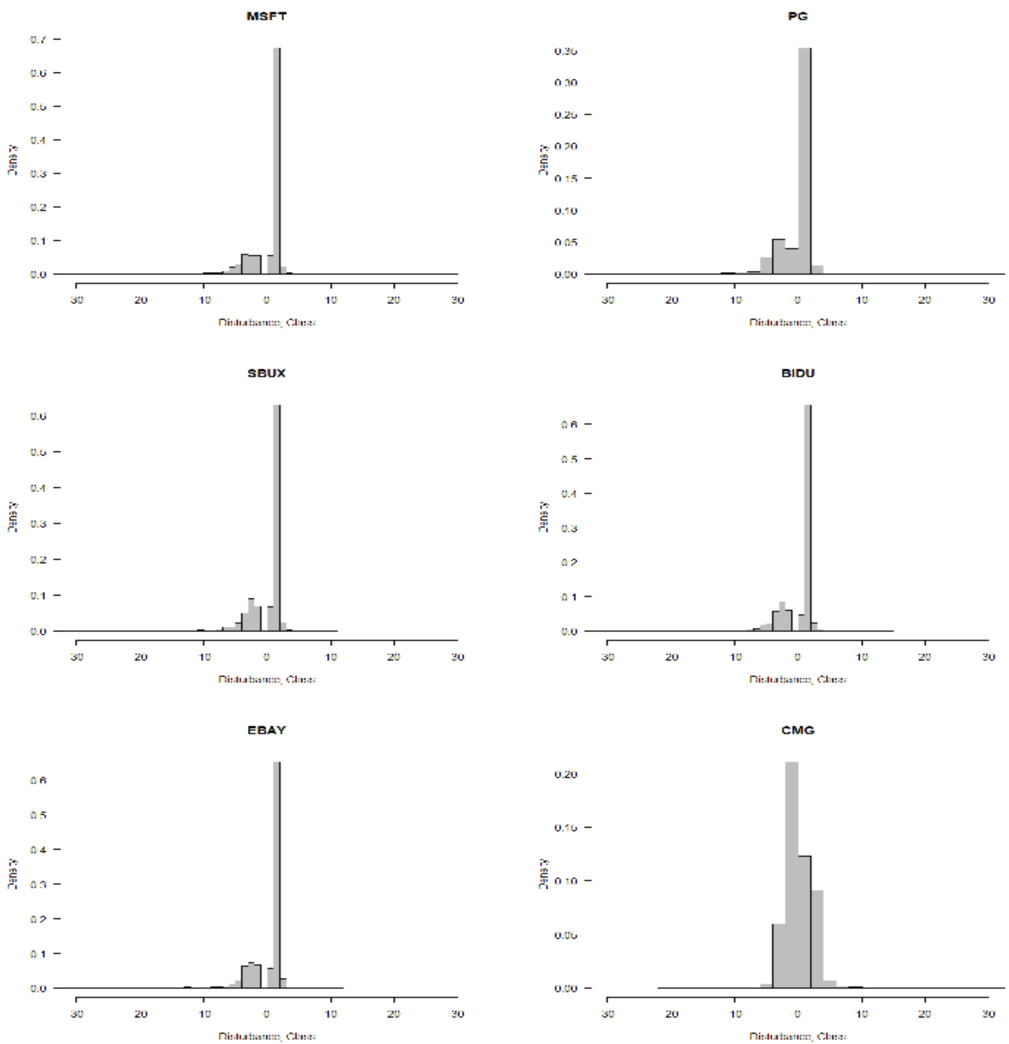



Figure 1.5: Histograms of disturbances as defined by Equation (8). Continued

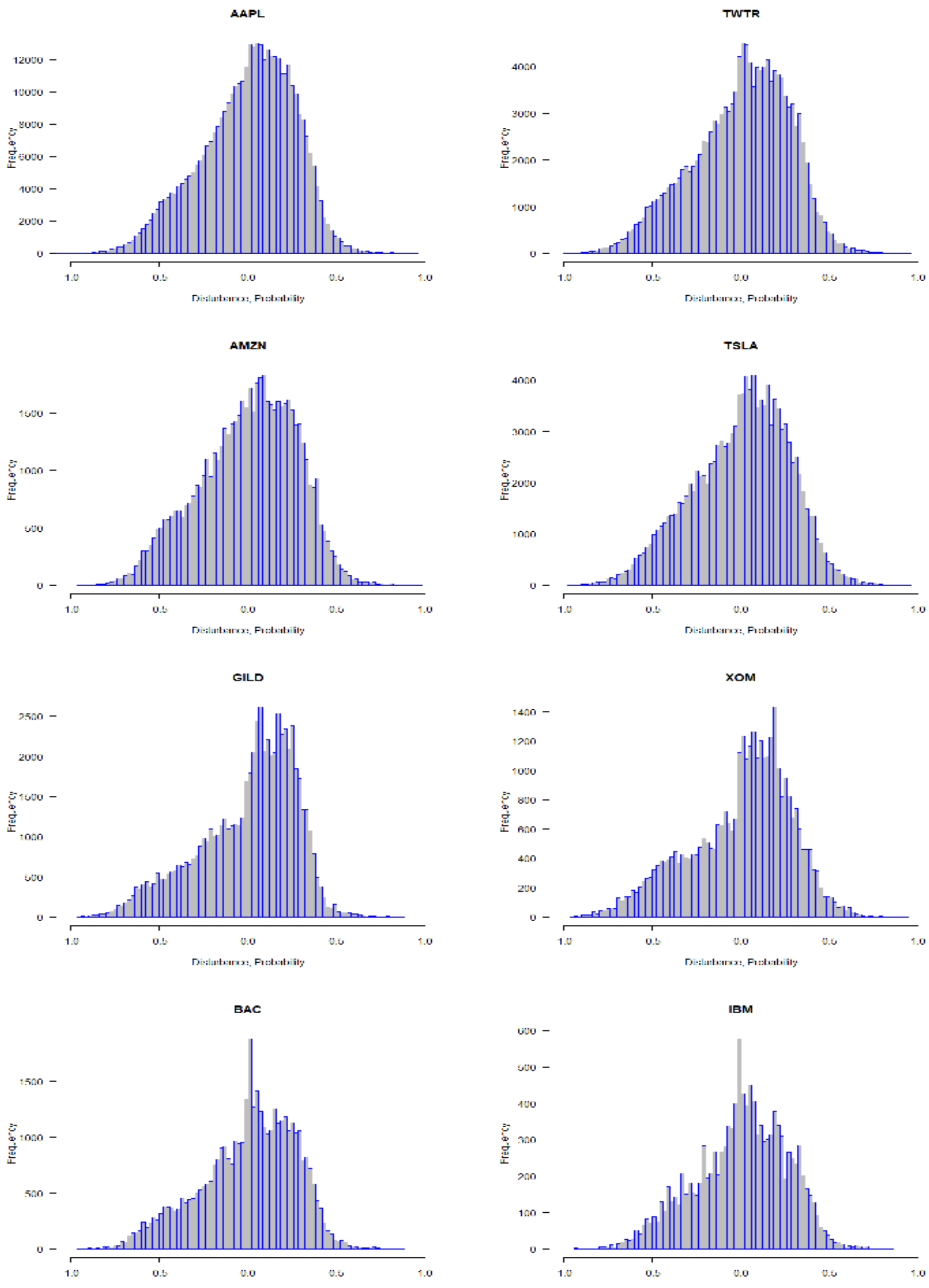



Figure 1.6: Histograms of disturbances as defined by Equation (13)

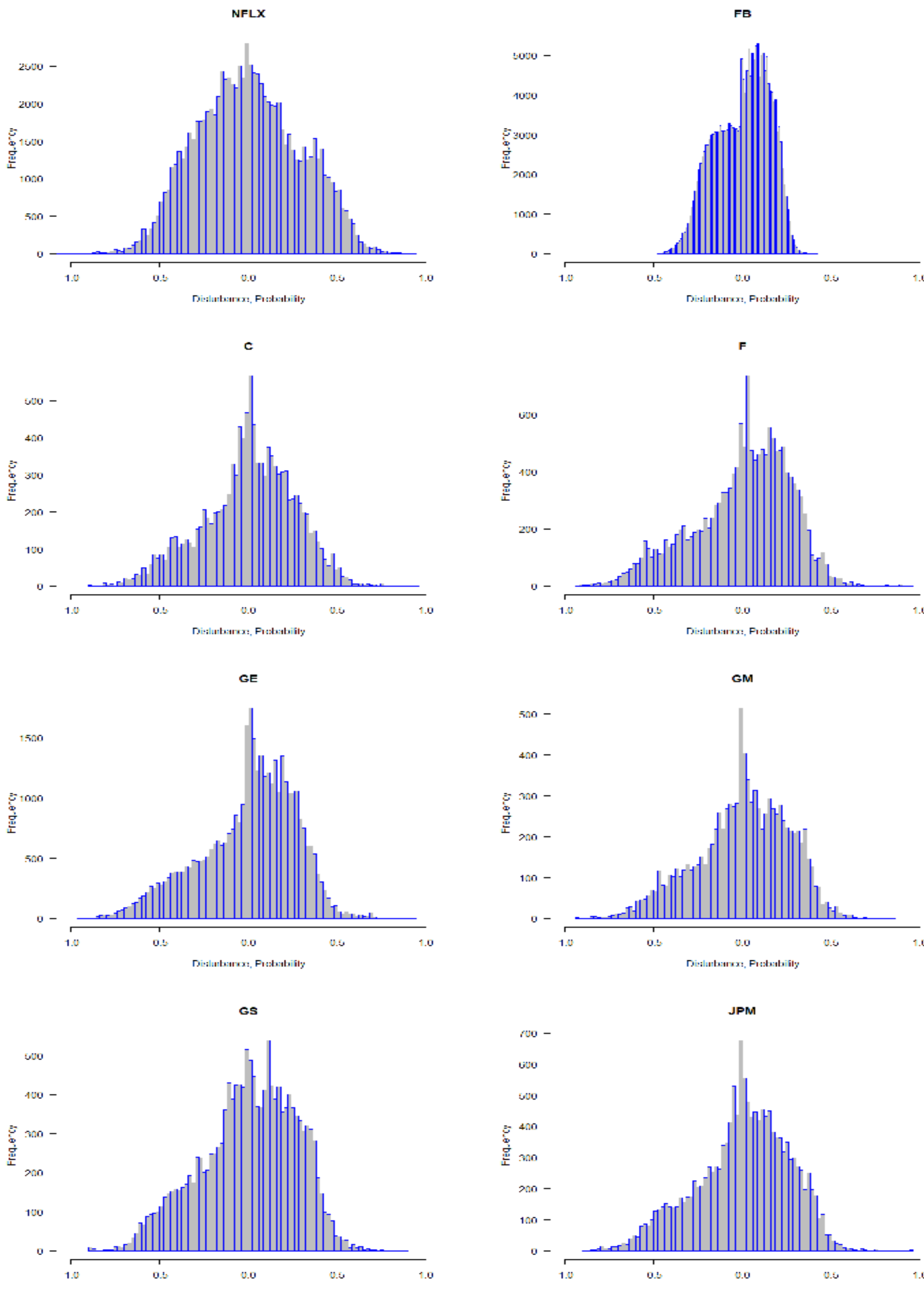



Figure 1.7: Histograms of disturbances as defined by Equation (13). Continued

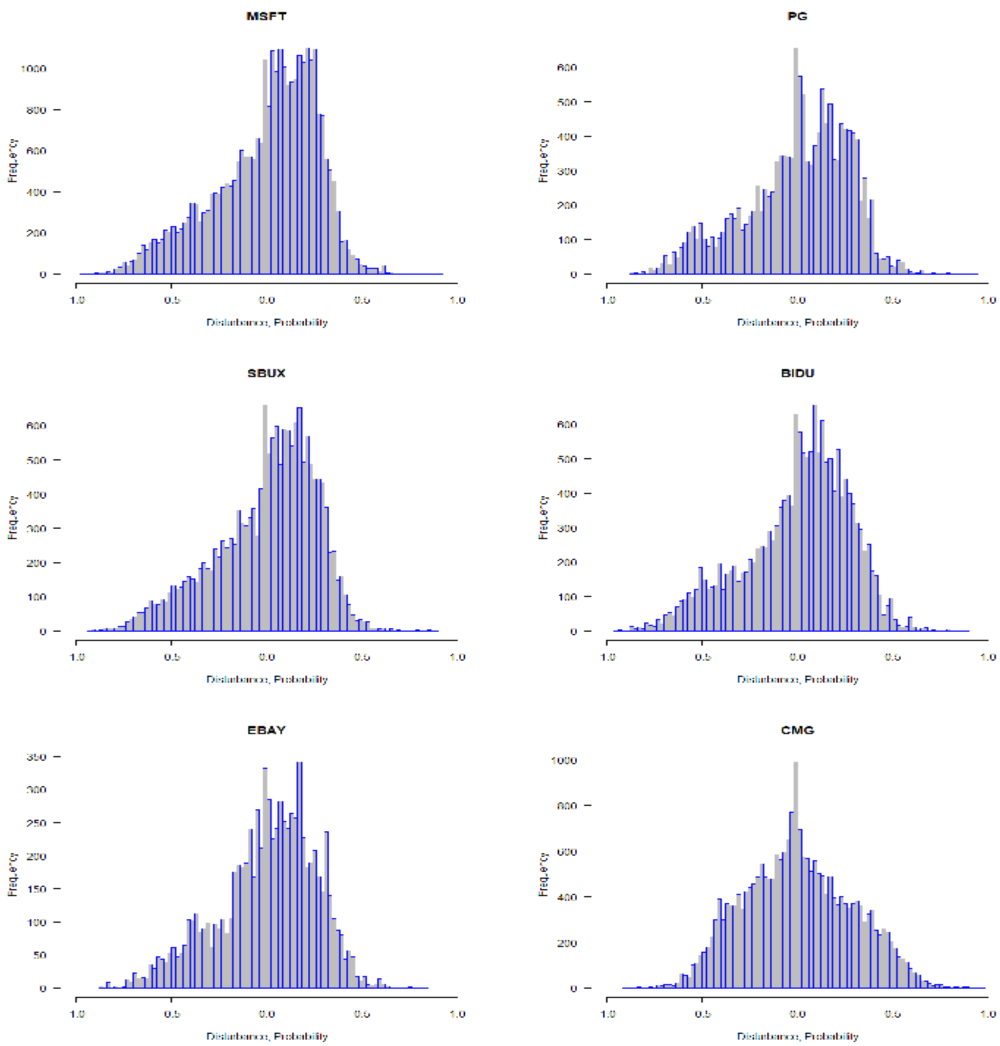
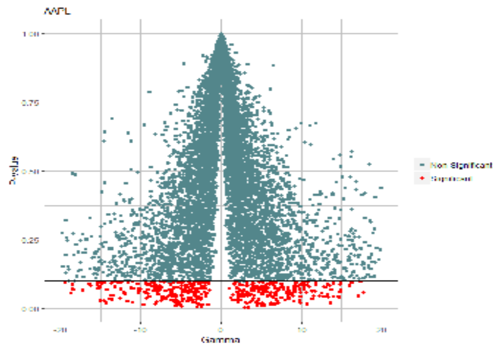
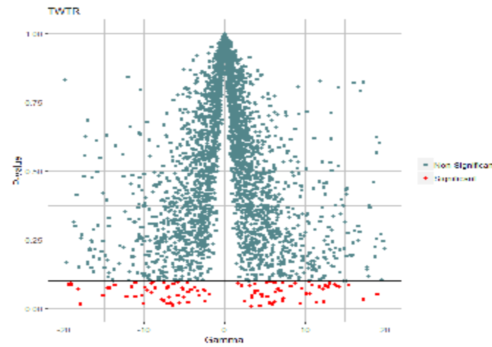


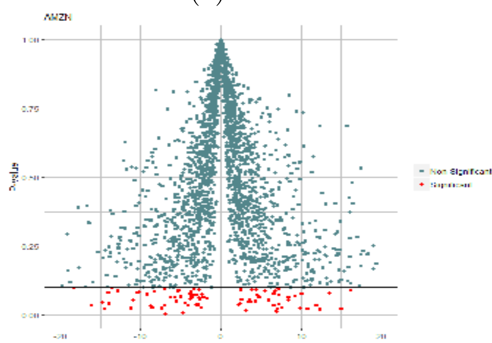

Figure 1.8: Histograms of disturbances as defined by Equation (13). Continued



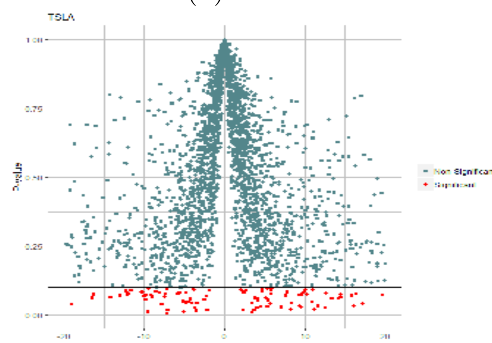
(a) AAPL



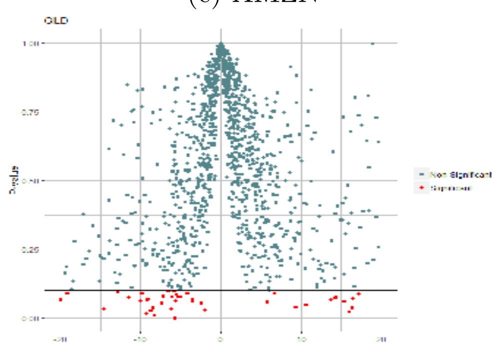
(b) TWTR



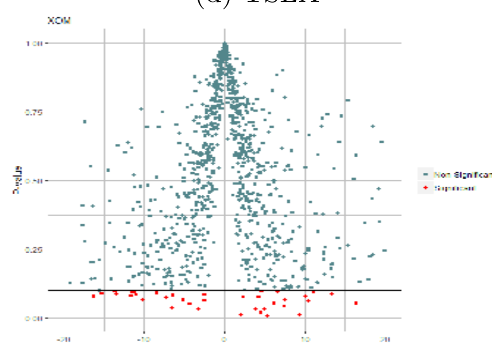
(c) AMZN



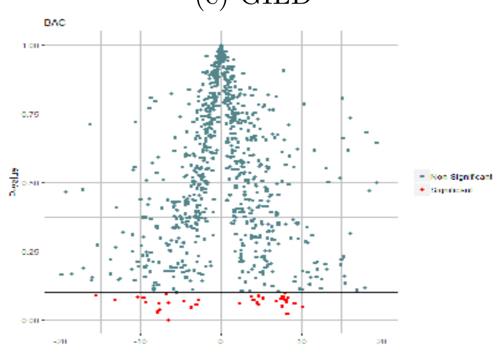
(d) TSLA



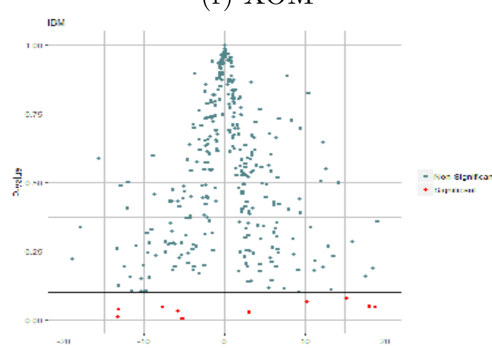
(e) GILD



(f) XOM

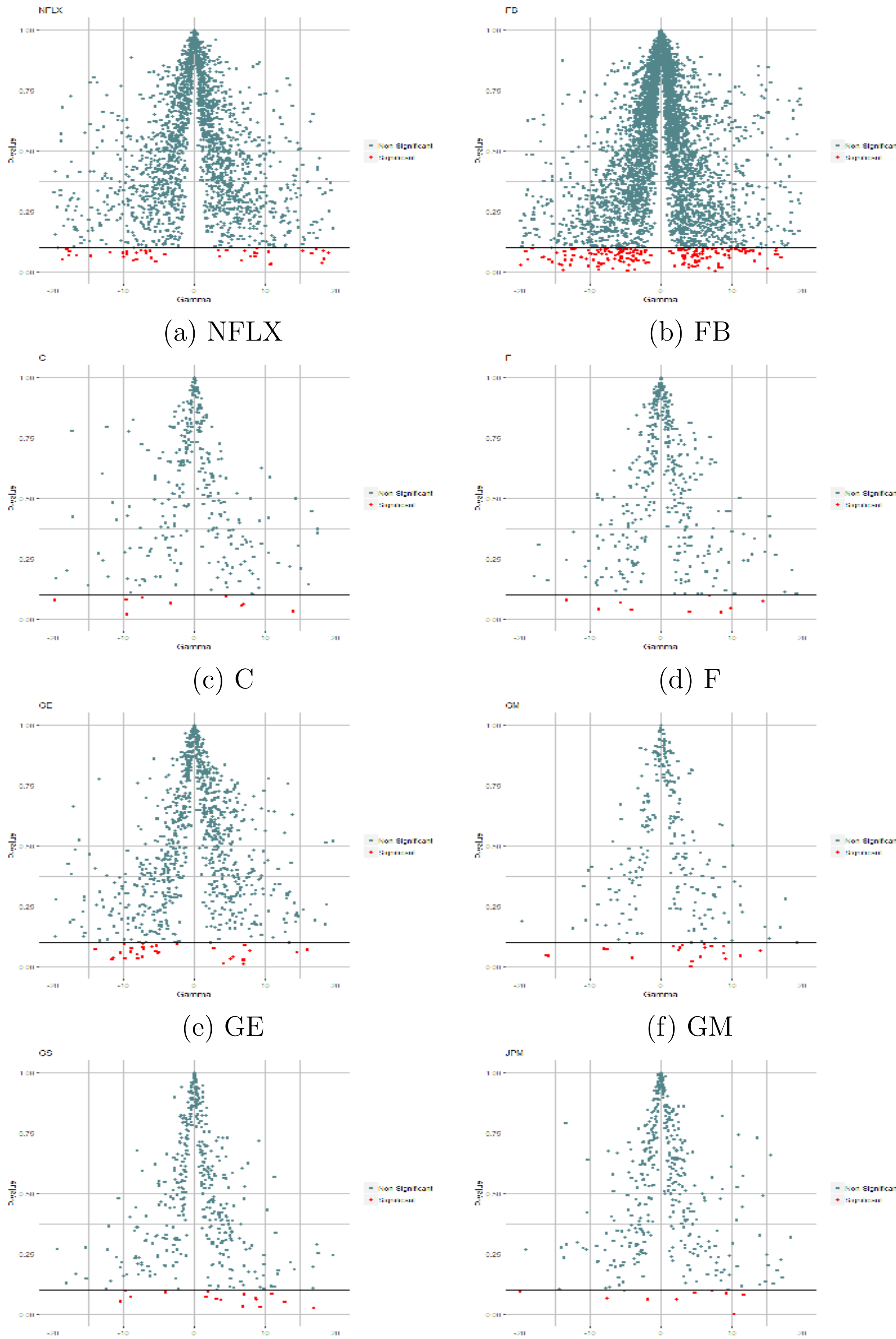


(g) BAC



(h) IBM

Figure 1.9: Value of Gamma against its P-value. Gammas as specified in Equation (8)




(a) NFLX

(b) FB

(c) C

(d) F

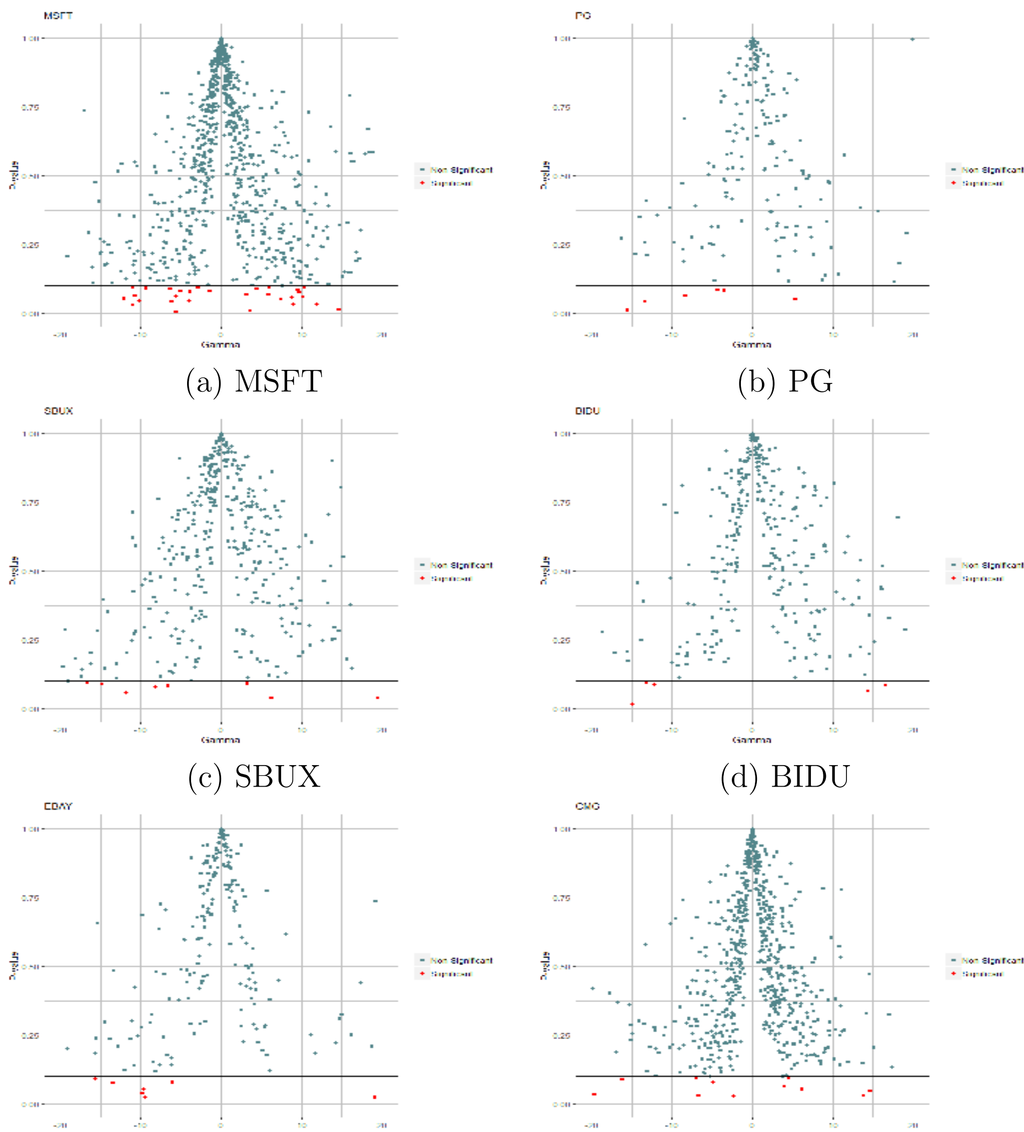
(e) GE

(f) GM

(g) GS

(h) JPM

Figure 1.10: Value of Gamma against its P-value. Gammas as specified in Equation (8). Continued

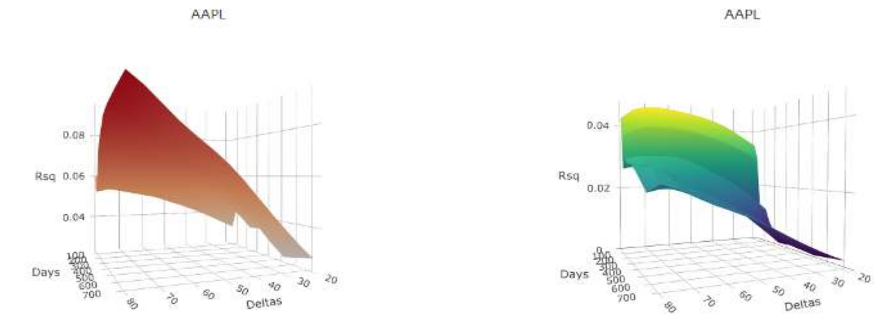



(a) MSFT (b) PG

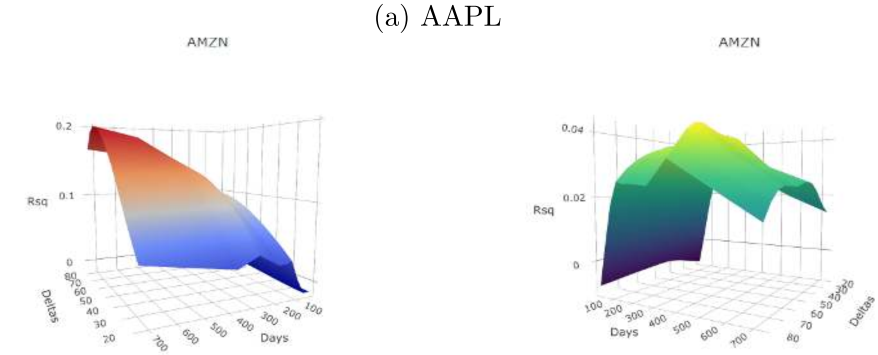
(c) SBUX (d) BIDU

(e) EBAY (f) CMG

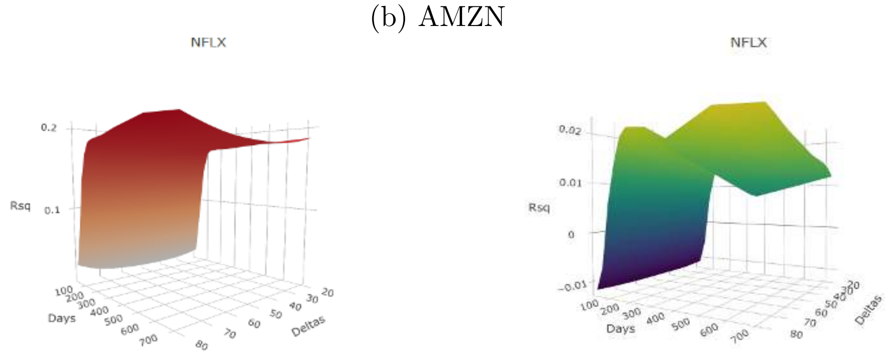
Figure 1.11: Value of Gamma against its P-value. Gammas as specified in Equation (8). Continued



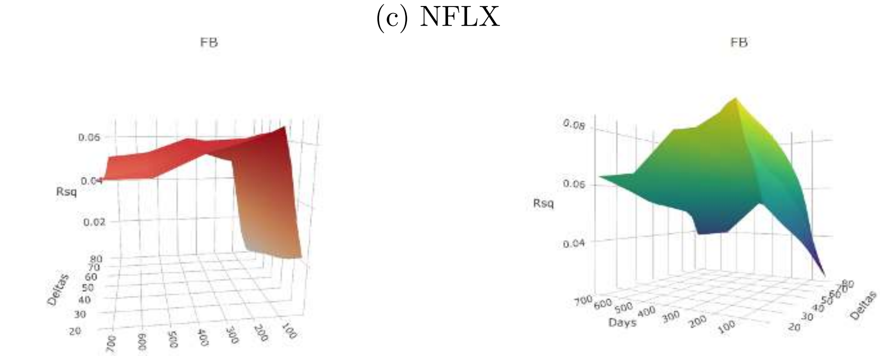

(a) AAPL




(b) AMZN

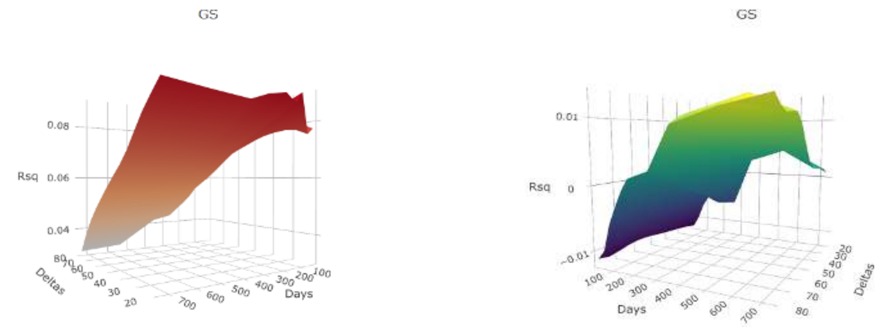



(c) NFLX

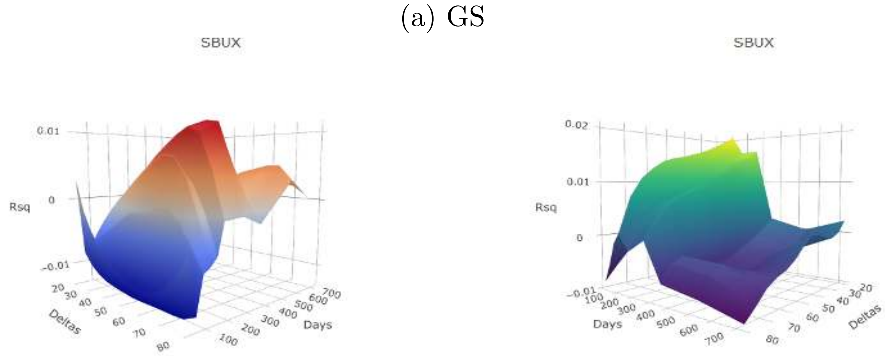



(d) FB

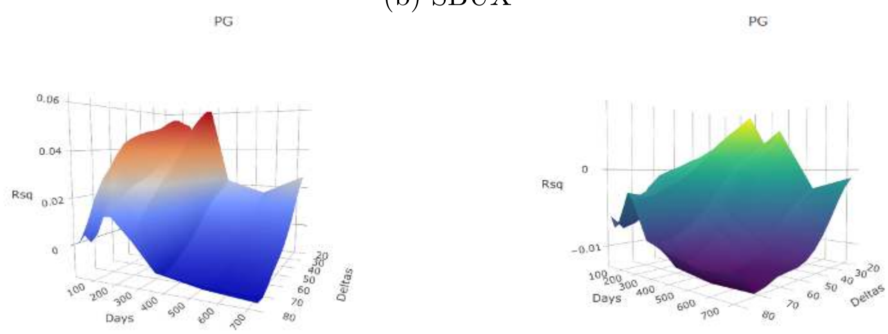
Figure 1.12: Portion of variation in the implied volatility (for a given pair of days to maturity and delta) explained by lags of $\rho$. Left panels represent $\rho$ obtained from non-directional W's, right panels represent directional W's



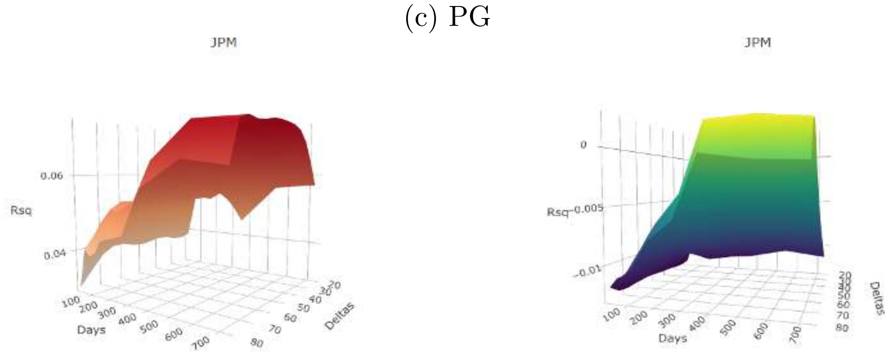

(a) GS




(b) SBUX

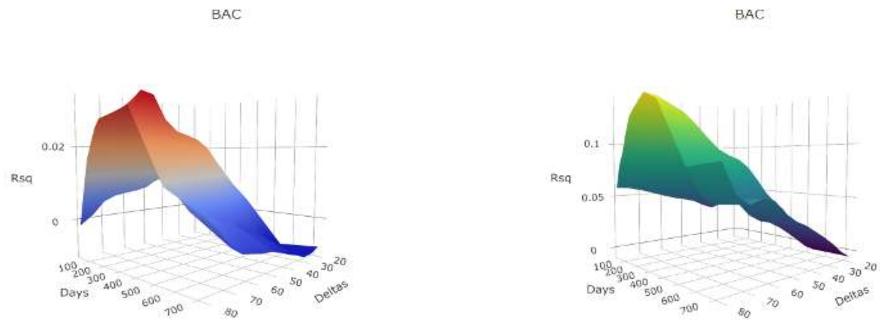



(c) PG

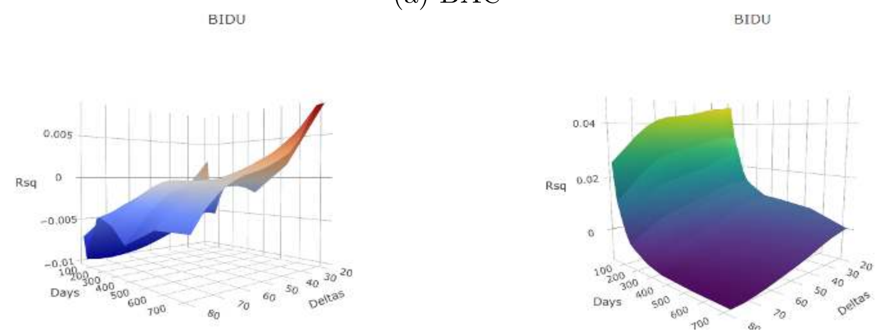



(a) JPM

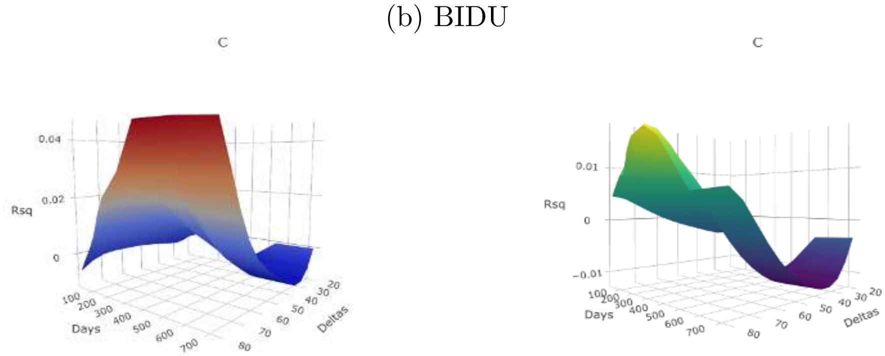
Figure 1.13: Portion of variation in the implied volatility (for a given pair of days to maturity and delta) explained by lags of $\rho$. Left panels represent $\rho$ obtained from non-directional W's, right panels represent directional W's. Continued



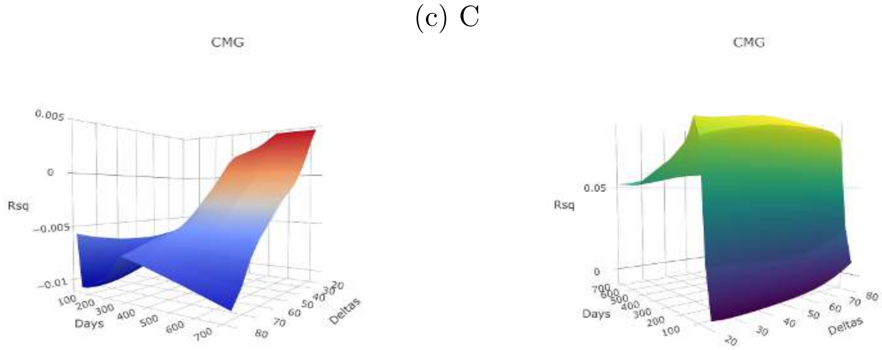

(a) BAC




(b) BIDU

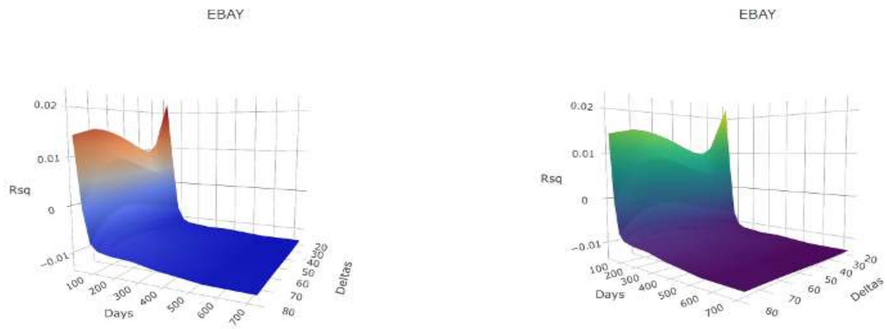



(c) C

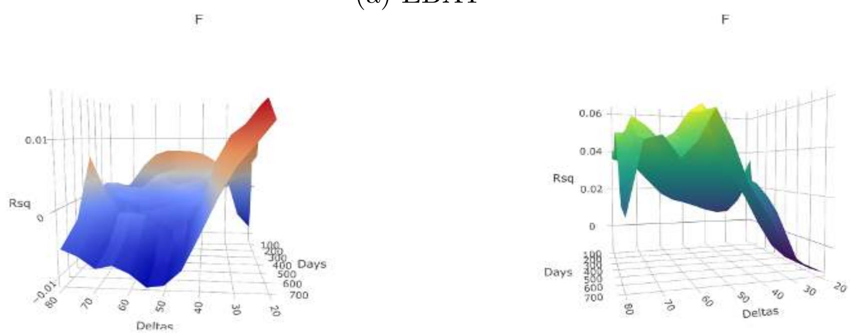



(d) CMG

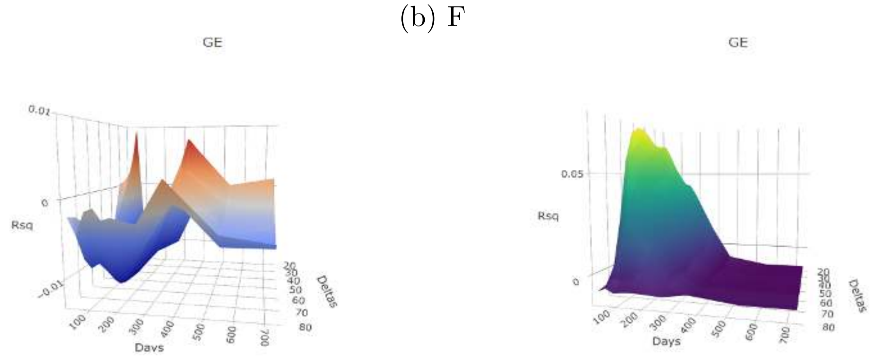
Figure 1.14: Portion of variation in the implied volatility (for a given pair of days to maturity and delta) explained by lags of $\rho$. Left panels represent $\rho$ obtained from non-directional W's, right panels represent directional W's. Continued



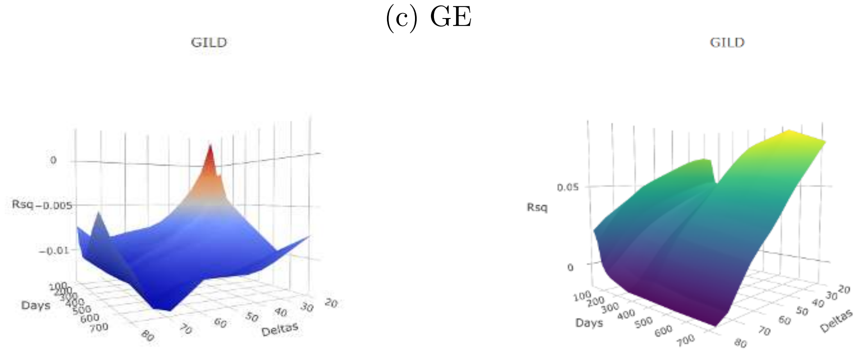

(a) EBAY




(b) F

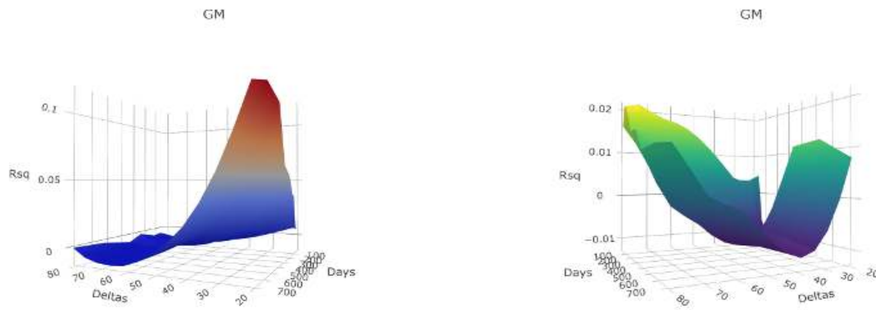



(c) GE

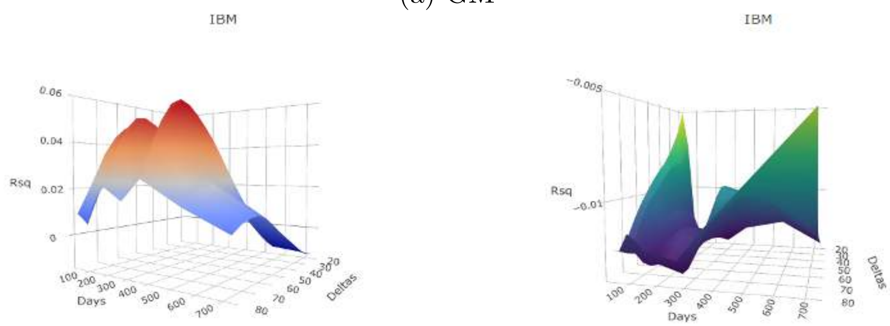



(d) GILD

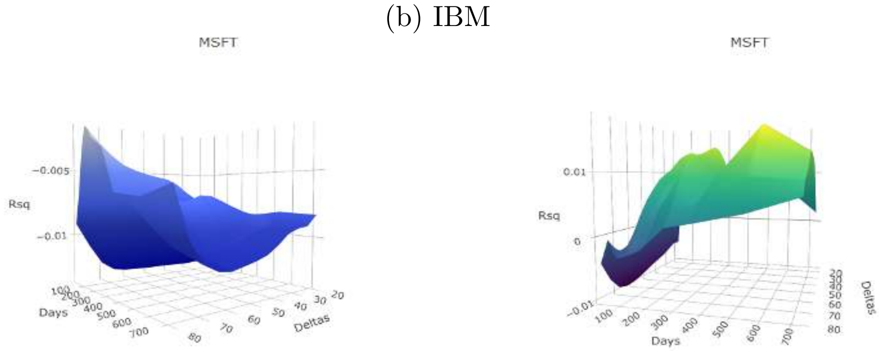
Figure 1.15: Portion of variation in the implied volatility (for a given pair of days to maturity and delta) explained by lags of $\rho$. Left panels represent $\rho$ obtained from non-directional W's, right panels represent directional W's. Continued



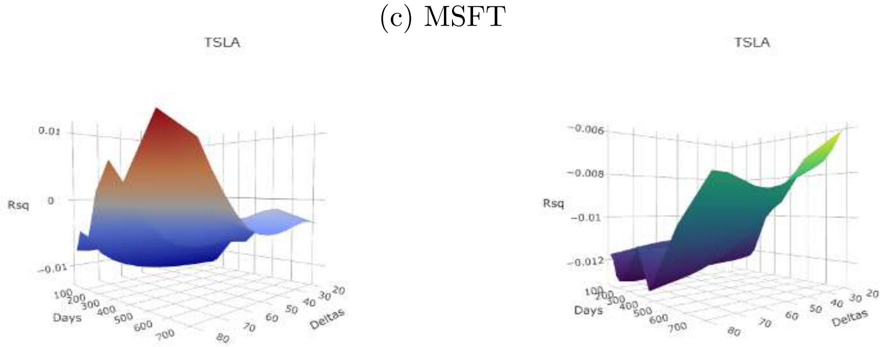

(a) GM




(b) IBM

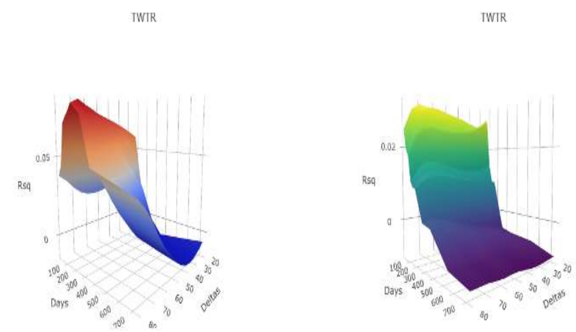



(c) MSFT

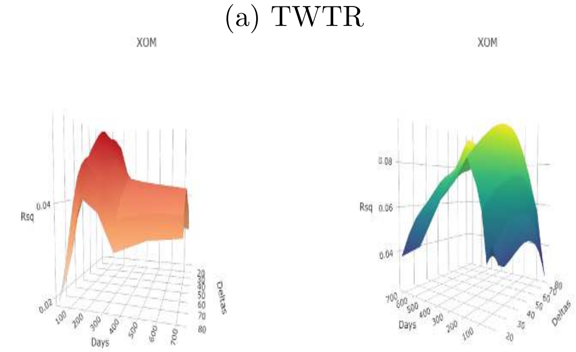



(d) TSLA

Figure 1.16: Portion of variation in the implied volatility (for a given pair of days to maturity and delta) explained by lags of $\rho$. Left panels represent $\rho$ obtained from non-directional W's, right panels represent directional W's. Continued

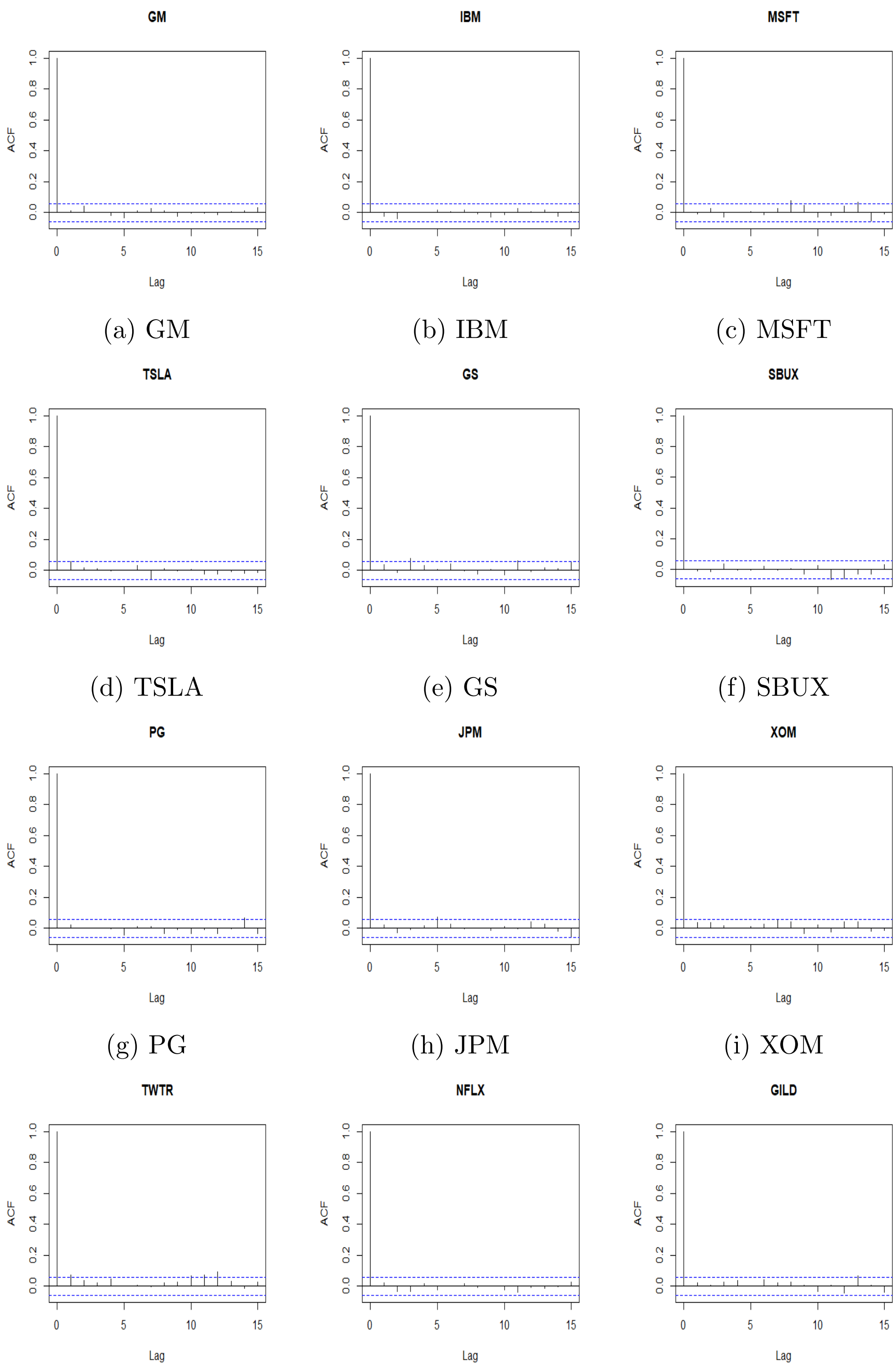



(a) TWTR




(b) XOM

Figure 1.17: Portion of variation in the implied volatility (for a given pair of days to maturity and delta) explained by lags of $\rho$. Left panels represent $\rho$ obtained from non-directional W's, right panels represent directional W's. Continued



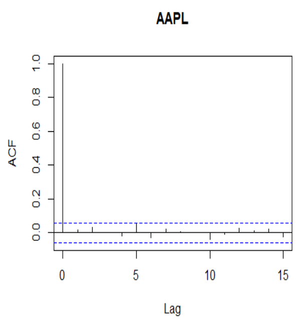

(a) GM (b) IBM (c) MSFT

(d) TSLA (e) GS (f) SBUX

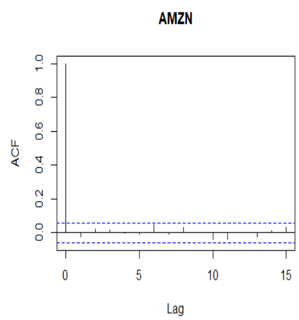
(g) PG (h) JPM (i) XOM

(j) TWTR (k) NFLX (l) GILD

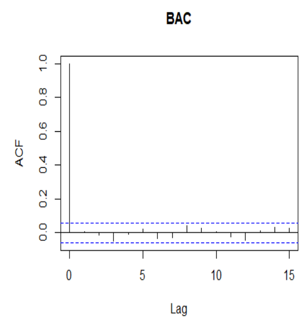
Figure 1.18: ACF plots for the News Polarity for each stock.



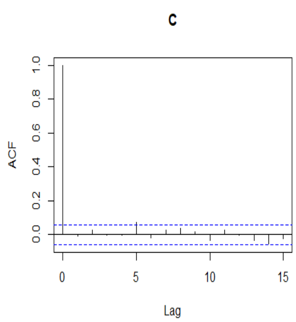

(a) AAPL



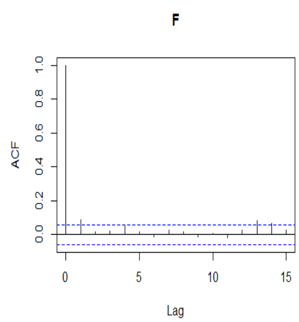

(b) AMZN



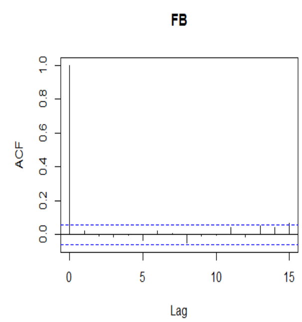

(c) BAC



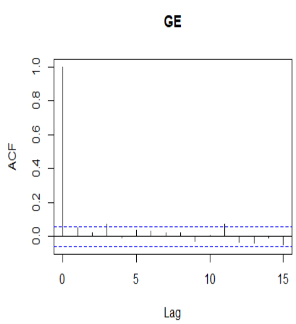

(d) C



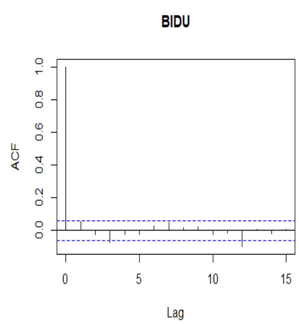

(e) F



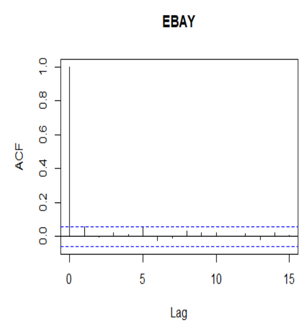

(f) FB



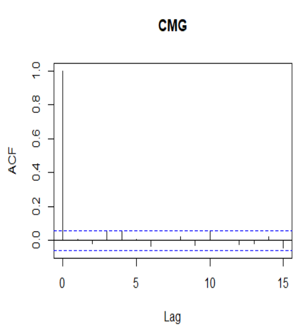

(g) GE




(h) BIDU

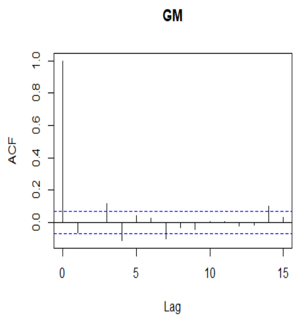



(i) EBAY

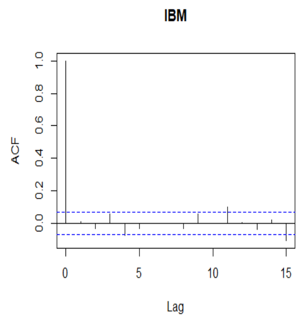



(j) CMG

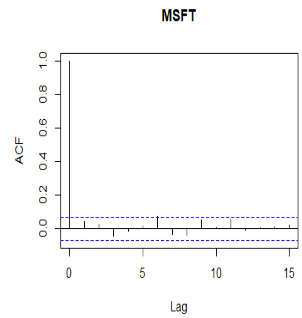
Figure 1.19: ACF plots for the News Polarity for each stock. Continued



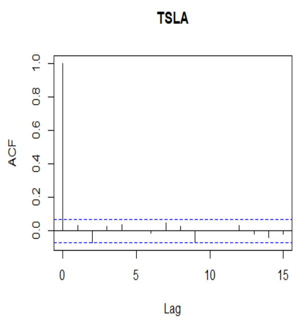

(a) GM



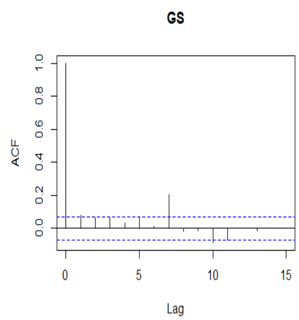

(b) IBM



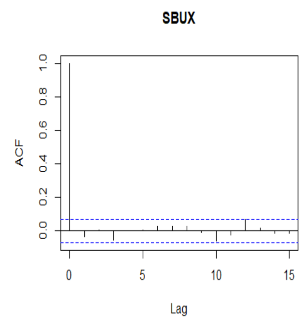

(c) MSFT



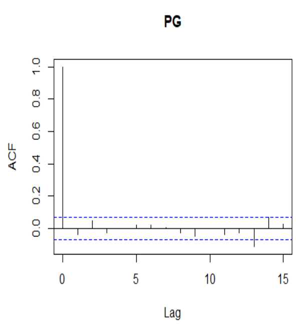

(d) TSLA



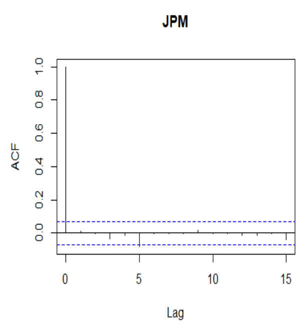

(e) GS



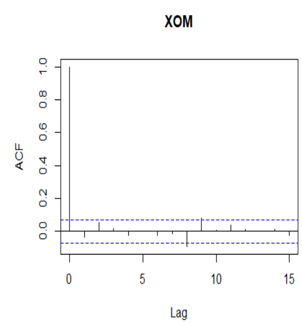

(f) SBUX



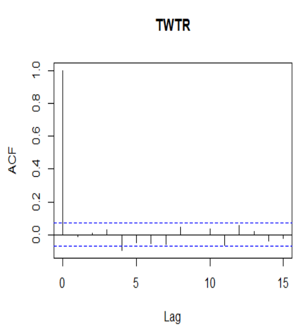

(g) PG



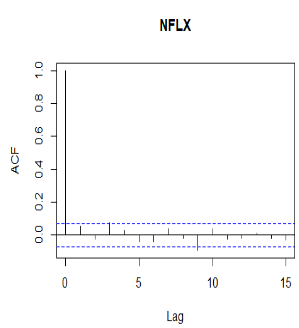

(h) JPM



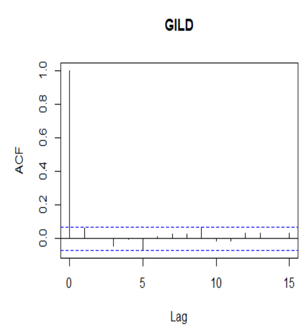

(i) XOM




(j) TWTR

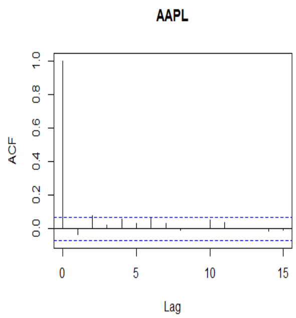



(k) NFLX

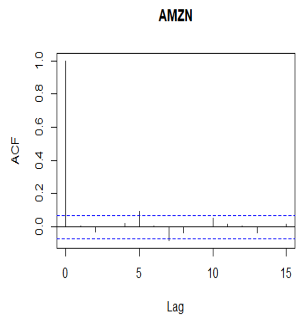



(l) GILD

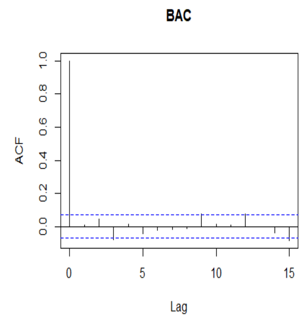
Figure 1.20: ACF plots for the herding rho for each stock.



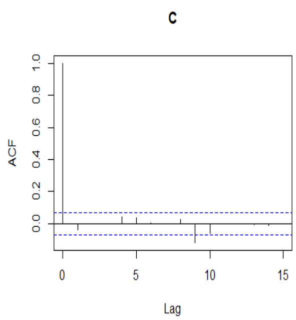

(a) AAPL



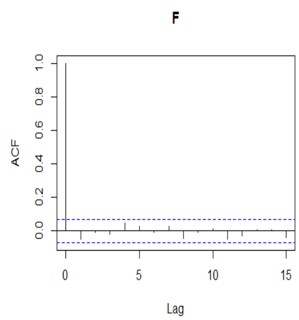

(b) AMZN



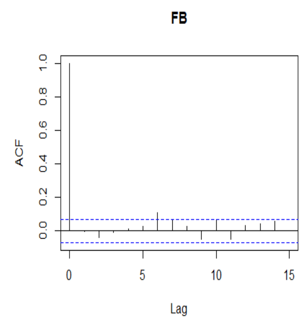

(c) BAC



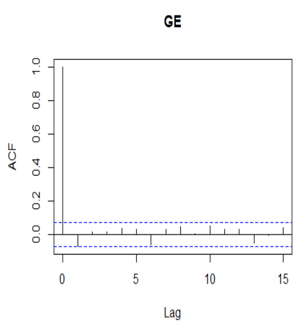

(d) C



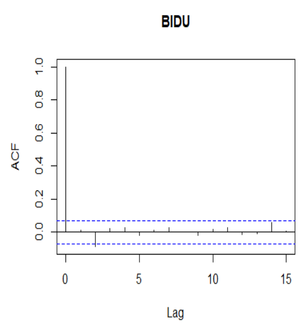

(e) F



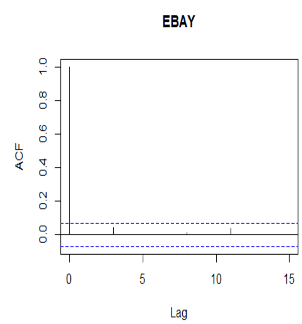

(f) FB



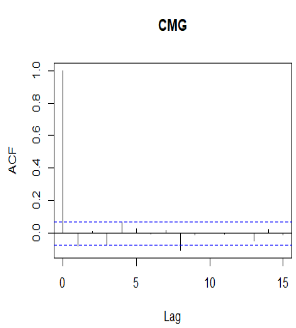

(g) GE




(h) BIDU

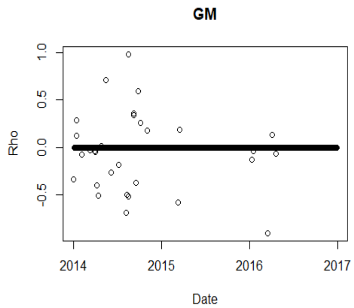



(i) EBAY

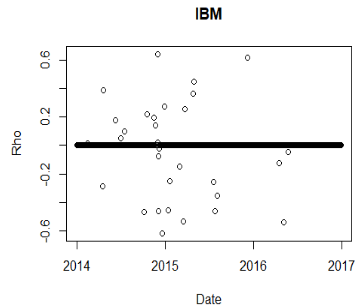



(j) CMG

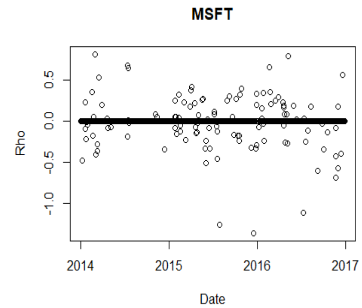
Figure 1.21: ACF plots for the herding rho for each stock. Continued



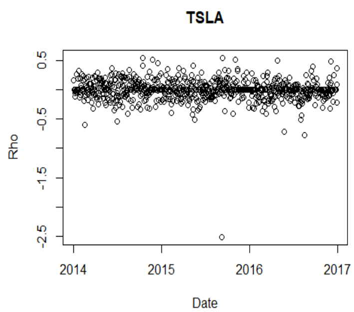

(a) GM




(b) IBM

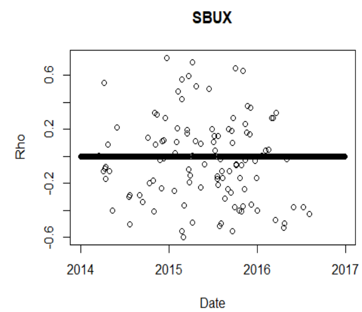



(c) MSFT

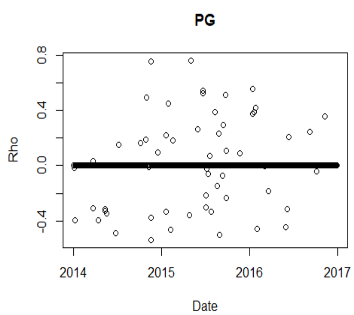



(d) TSLA

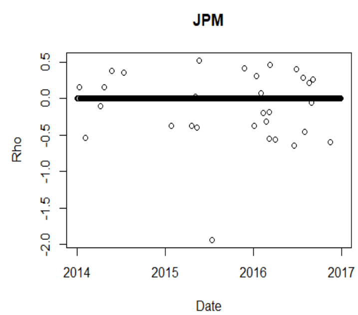

(e) GS



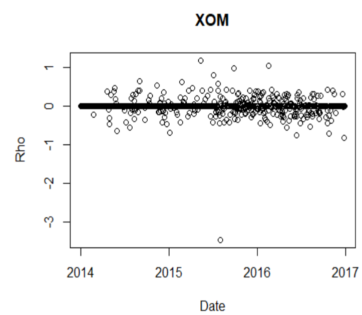

(f) SBUX



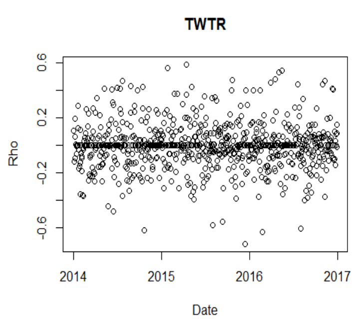

(g) PG



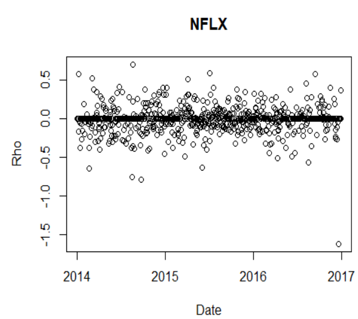

(h) JPM



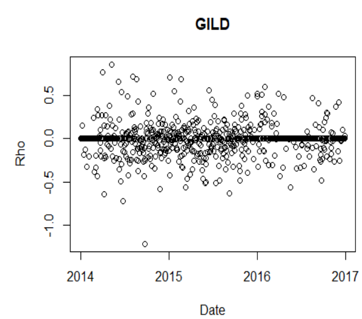

(i) XOM




(j) TWTR

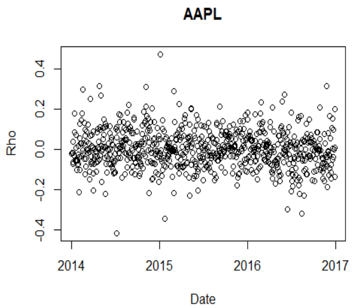



(k) NFLX

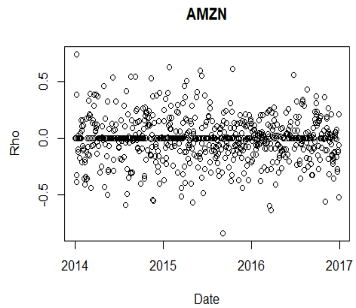



(l) GILD

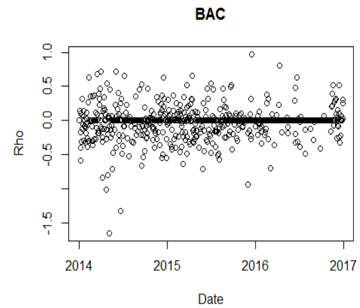
Figure 1.22: Plots of herding rho over time for each stock.



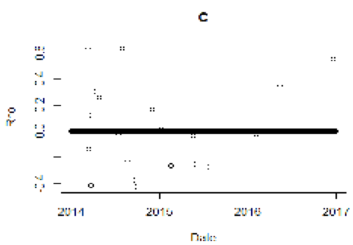

(a) AAPL



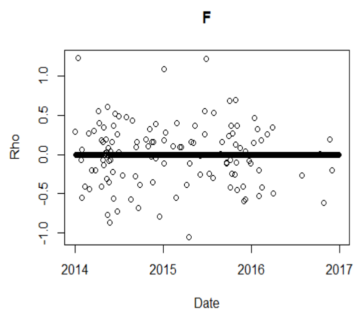

(b) AMZN



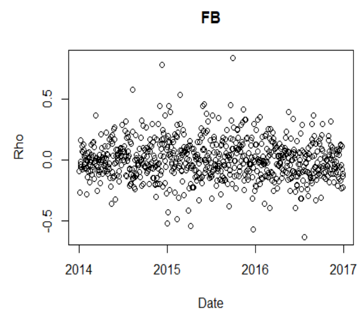

(c) BAC



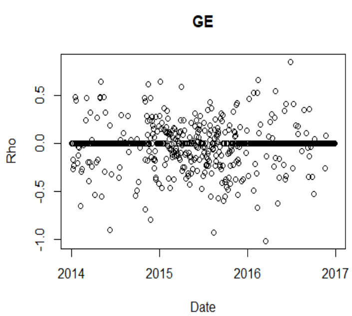

(d) C



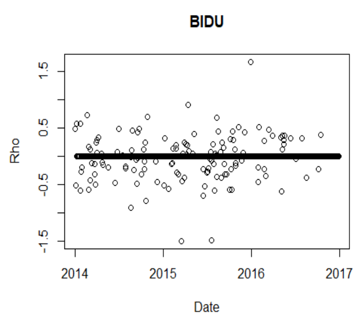

(e) F



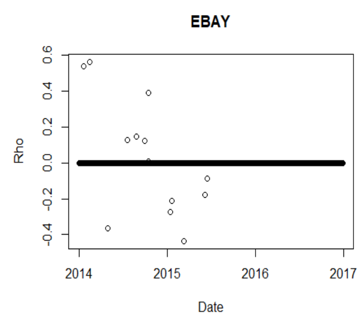

(f) FB



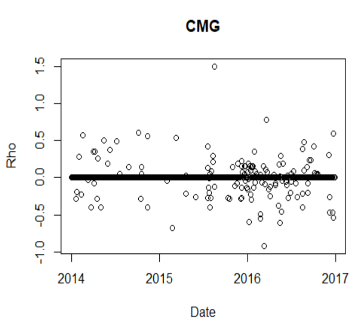

(g) GE




(h) BIDU

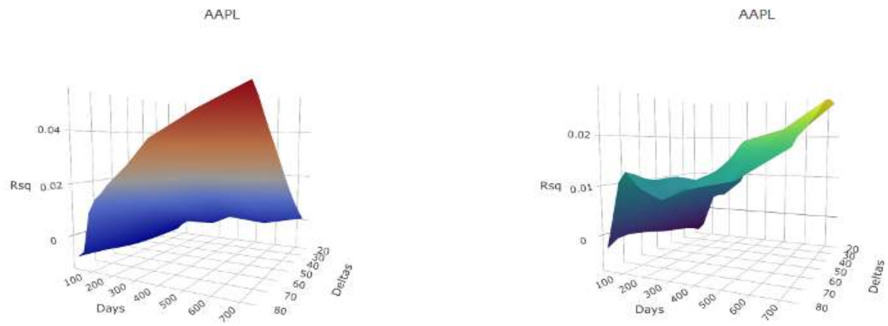



(i) EBAY

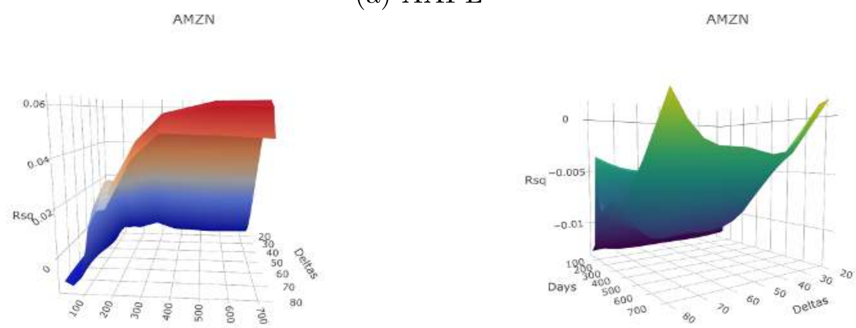



(j) CMG

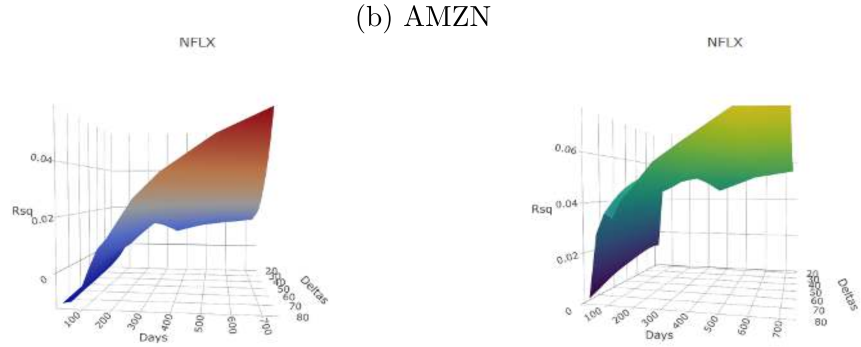
Figure 1.23: Plots of herding rho over time for each stock. Continued



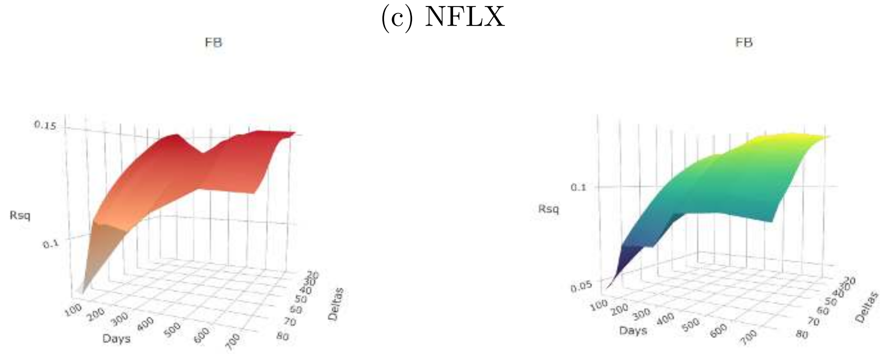

(a) AAPL




(b) AMZN

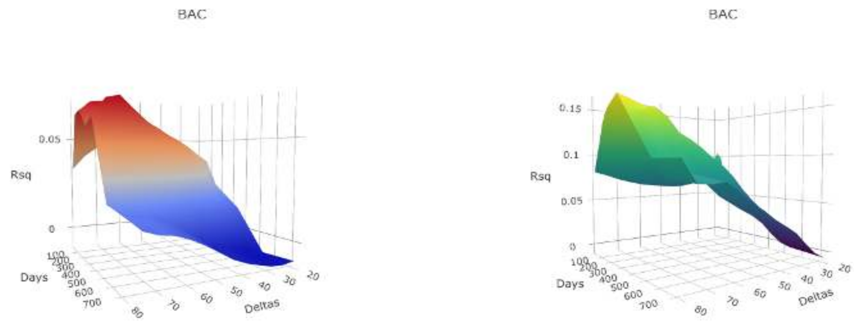



(c) NFLX

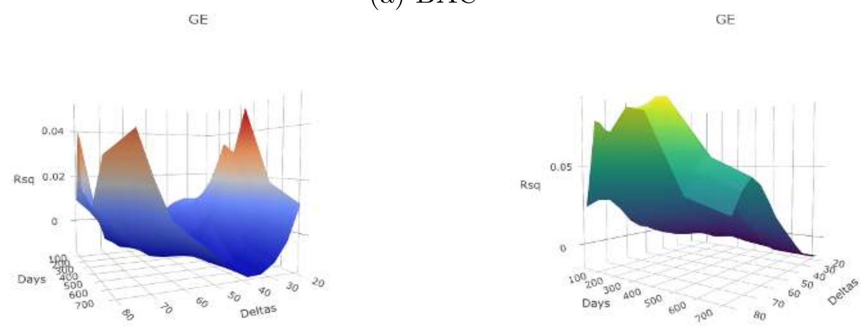



(d) FB

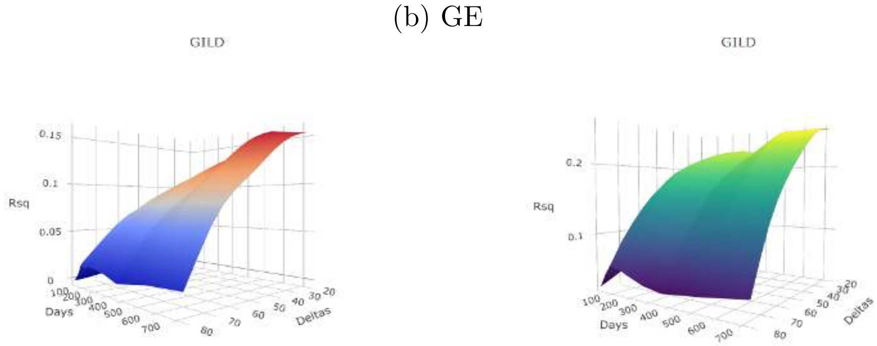
Figure 1.24: Portion of variation in the implied volatility (for a given pair of days to maturity and delta) explained by lags of $\rho$. Left panels represent $\rho$ obtained from non-directional W's, right panels represent directional W's. Decisions filtered through seven lags of news.



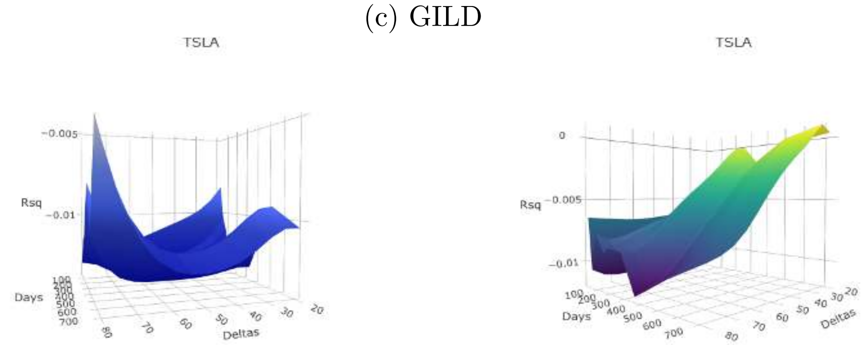

(a) BAC




(b) GE

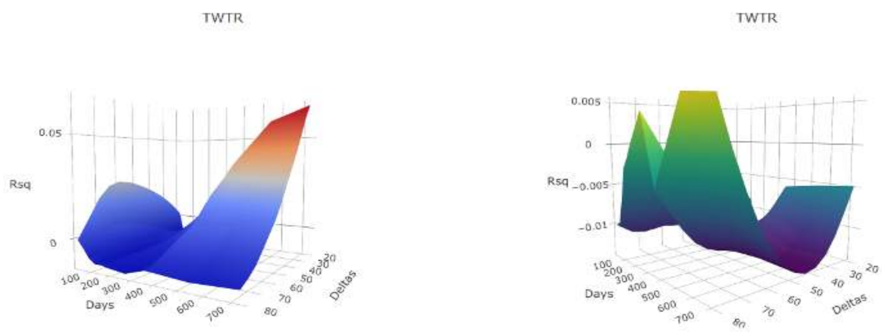



(c) GILD

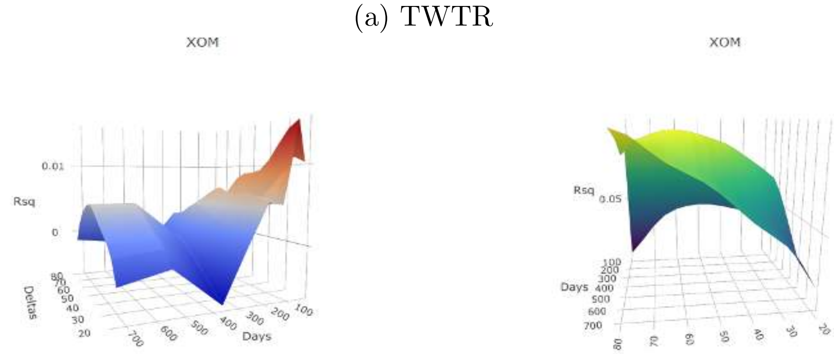



(d) TSLA

Figure 1.25: Portion of variation in the implied volatility (for a given pair of days to maturity and delta) explained by lags of $\rho$. Left panels represent $\rho$ obtained from non-directional W's, right panels represent directional W's. Decisions filtered through seven lags of news. Continued




(a) TWTR




(b) XOM

Figure 1.26: Portion of variation in the implied volatility (for a given pair of days to maturity and delta) explained by lags of $\rho$. Left panels represent $\rho$ obtained from non-directional W's, right panels represent directional W's. Decisions filtered through seven lags of news. Continued

# Chapter 2

# Banks under the Volcker Rule: Elastic Net Approach

## 2.1 Introduction

As a response to the financial crisis, the Dodd-Frank Wall Street Reform was signed into law on July 21, 2010. A part of it was a special provision that prohibits banks from the proprietary trading and investing in hedge funds and private equity – the so-called Volcker Rule.

Prop trading, as defined by the regulation, is "engaging as a principal [or main party to a transaction acting for his own account and risk] for trading account of the banking entity in any transaction to purchase or sell security, derivative, or any other financial instrument". Without restriction on such activity, large banking entities could allow themselves to take on increased risk in trading while still relying on commercial banks access to the Fed Banks Discount Window: availability of cheap fed funds makes elevated risk levels more tolerable, but at the expense of the US taxpayer (Chatterjee (2011)).

The rule that once was envisioned as a simple clear ban on prop trading, gained complexity while being worked on by five regulatory agencies. It is argued by some that it could deliver benefit of financial stability if implemented properly (Baily et al. (2017)); however, the cost of implementation would be high. For instance, JPMorgan's estimate of only early compliance cost was 400-600 million annually.

In addition to direct cost of implementation, one would expect to observe an indirect cost of forgone profits due to the ban of certain trading activities. Yet, for most banks that made changes in order to comply with Volcker, earnings reports did not show much difference compared to pre-Volcker times. The percentage of total revenue these banks generated from trading also remained quite stable over time. There were two possible explanations for such observations floating in the literature. The first pointed out multiple methodological flaws of the Rule (Chatterjee (2011)): a return to Glass-Steagalls restrictions is problematic after the Gramm-Leach-Bliley Act – once allowed by the latter, the financial combinations of commercial and investment banks became a well-tied structure that is vital to todays financial markets. Thus, the Volcker Rule, focuses not on structure but rather on transaction — it bans prop trading for deposit-taking banks, but exempts market-making and hedging activities. In particular, the distinction between speculative trade and hedge oftentimes is hard to make.

Secondly, banned from one type of trading, banks are trying to make up their revenue by engaging more in the type of trading that is still allowed, as some speculate (Chung et al. (2019)). Because of increased engagement in such seemingly safer transactions, an aggregate level of risk would not decrease (and potentially could even get higher), defeating the whole purpose of the Volcker Rule.

In this study, I test such spillover effect. I find no impact of the Rule on measures of profitability (RoE), which is consistent with the previous studies. A potential explanation here is that the variable possesses an exogeneity characteristic: it is rather a measure that banks aim to attain and relocate their resources in such an order so that the bar is met.

I speculate, however, that, following the enactment of the Volcker Rule, there is a change in the banking practices. To test this conjecture, I consider banks' non-interest revenue generated by the activities similar to those banned under the Volcker Rule. In addition, I propose a new way of identification: I collect banks' feedback about the Rule's potential and current impact on their operations and financial results from the annual 10-K reports. This approach allows for more accuracy in separation between those banks that are subject to the law, and the banks in the controlling group – a feature much needed in this policy evaluation question. I use a synthetic control method in the form of an elastic net approach, which allows me to relax some assumptions posed by the traditional approach and to tailor the method to the peculiarities of the banking business.

I find that the notional values of derivatives held for hedging increase in response to the Rule for most of the banks with the result being statistically significant for the majority of the affected banks. The percentage of the revenues brought by investment banking advisory, brokerage, and underwriting also shows a small increase, although, the result is not statistically significant; similarly, trading assets and riskiness of banking activity (as proxied by z-score) decrease in post-Final Rule era, and the result is significant for for the few last observed months in the testing window. Per recent developments in the regulation, available for sale securities were also considered. No statistically significant change was captured.

## 2.2 Background

In 1933, the Congress passed the Glass-Steagal Act in order to restore the confidence in the banking system. In a prelude to the Act, banks were allowed to underwrite and deal in securities and the decade of 1920s was extremely prosperous up until 1929 when the "boom" turned into a "bust", with worried depositors rushing to the banks in hope to withdraw their funds. Such bank runs have resulted in the Emergency Banking Act, also named

Glass-Steagal Act, separating commercial banks from the investment banks.

Along with the separation of banking activities, the Act specified the creation of the FDIC, which started insuring bank deposits with a pool of money gathered from the banks. The deposit insurance had in its key function to promote confidence in the banking system and prevent it from the aforementioned "bank runs." Such guarantees, however, potentially invited excessive risk-taking – a so-called moral hazard effect: lacking control by the depositors that are ready to withdraw their funds in case of bank being unsafe, the bank is more prone to get involved in a risky action. The risk, in turn, was transferred to the public.

In 1999 the Glass-Steagal Act was repealed by Gramm-Leach-Bliley Act with the purpose to modernize financial system. The Act allowed for the creation of financial behemoths that housed the deposit banks and often risky investment business under a single roof. It was passed with an enormous fanfare; however, as Macey (2000) notes, at the time of its passage Glass-Steagal Act was already a dead letter: "the investment banks had conspired with compliant regulators to punch giant holes in the statutory restrictions on combining commercial banking and investment banking."

By that time, the FDIC insurance limit grew up to \$100,000. In 2008, Congress approved a temporary increase in the limit to \$250,000 but the Dodd-Frank Reform and Consumer Protection Act of 2010 made this number permanent. The moral hazard question was raised once again with the suggestion of possible solution to the problem. During his testimony before the Senate Banking Committee, Paul Volcker mentioned:

> *The basic point is that there has been, and remains, a strong public interest in providing a safety net – in particular, deposit insurance and the provision of liquidity in emergencies – for commercial banks carrying out essential services. There is not, however, a similar rationale for public funds — taxpayer funds — protecting and supporting essentially proprietary and speculative activities. Hedge funds, private equity funds, and trading activities unrelated to customer needs and*

> *continuing banking relationships should stand on their own, without the subsidies implied by public support for depository institutions.*

The Volcker proposal to limit banks proprietary trading and fund activity was signed into law in June 2010 as Section 619 of the Dodd-Frank Act. The text of it, however, required extensive agency definition and rule-making. In particular, the first part of it prohibits banking entities from engaging as principal in proprietary trading for the purpose of selling financial instruments in the near term or otherwise with the intent to resell in order to profit from short-term price movements. The prohibition is subject to numerous exemptions, including trading in U.S. government, agency and municipal obligations, underwriting and market making-related activities, risk-mitigating hedging activities, trading on behalf of customers, trading for the general account of insurance companies, and foreign trading by non-U.S. banking entities. Oftentimes, such activities as market making, hedging, and underwriting evidence quite similar characteristics to the prop trading making it difficult to draw a distinct line between them. For instance, accumulation of the inventory in anticipation of the customer demand might resemble prop-trading; banks also might engage in prop trading through incomplete or inconsistent hedging strategies, as noted in the FSOC study (FSOC (2011)). Broadly determined restrictions may have a deterring effect on the permitted activities; however, the loose definition of prohibited area might invite an opportunity for the camouflaged prop trading.

The second part of the Volcker Rule bans the bank holding companies from investing in or sponsoring private equity or hedge funds (so-called covered funds), again subject to a list of exemptions. Funds that are not regarded as covered include foreign public funds, wholly owned subsidiaries, joint ventures, acquisition vehicles, securitization-related vehicles, mutual funds, exchange-traded funds, etc. In addition, the banks are permitted to take part is so-called "de-minimis investment" in which the bank might provide up to 100% of the seed capital or make investment in a fund as long as the amount of it is no more than 3%

of the total ownership interest within a year from transaction. Such investment must also be immaterial to the bank and should constitute no more than 3% of Tier 1 Capital of the banking entity.

With regard to the latter portion of the Rule, the federal regulators and others expressed worries about the potential market disruptions caused by banks dumping the covered funds into the marketplace, as well as about the burdens the banks would experience in case of liquidation of already held investments. As a result, the Volcker Rule implementation was prolonged by the extensions given to banks on their way to compliance with the legislation.

To this day, Volcker Rule remains to be a hotly debated topic in the literature. A number of studies point out its potential unintended negative impact on the bond market liquidity (Duffie (2012); Bao et al. (2016); Whitehead (2011); Chow and Surti (2011)); some suggest its complete ineffectiveness in today's financial markets structure (Chatterjee (2011); Chung et al. (2019)); some empirical studies find conflicting evidence on the market liquidity post crisis and during the Volcker Rule enactment (Trebbi and Xiao (2015); Bessembinder et al. (2018); Dick-Nielsen and Rossi (2016)). Bao et al. (2016) notes that such inconsistencies in results might stem from the differences in periods examined in the studies – some authors focus on pre-Volcker events and discuss their results in light of anticipation of new regulation rather than assessing implementation of the Rule and its impact on market conditions. The latter would require the extension of the time window to, preferably, today's date.

Few studies bring their analysis to a firm (bank) level, looking into the impact of Volcker on banks default likelihood and profits. Chung et al. (2019) find that the impact on default probability could be unfavorable; profitability, however, is expected to decrease, based on their simulation results. In their empirical exercise, the authors find no evidence of such decrease in profits, and explain this inconsistency by their sample limitation: their sample consists of banks with higher banking profitability than their trading activity. Moreover, the Rule decreases the trading book and raises the illiquid banking book portfolio, which yields even higher earnings at the expense of higher risk.

Other studies suggest that the Volcker Rule could reduce banks' risk management capabilities and the services that they offer (Thakor and E. (2012)). Madura and Premti (2014) find a negative valuation effect for money center banks following the Rule's announcement. They also note a risk reduction as a response to a first event signaling development of Volcker and attribute such change to the perception of the banking industry: the Rule boosts confidence of the investor in the risk-taking behavior of a bank. In contrast to these results, Keppo and Korte (2016) find no decrease in default probabilities for affected banks in post-Volcker times: the volatility of trading returns remained unchanged and volatility of banking activities decreased, suggesting that risk-taking has not moved to the banking book.

## 2.3 Model and inference

The standard approach in the policy evaluation literature follows Mill's Method of Difference (Mill (2011)), which specifies:

> *If an instance in which the phenomenon under investigation occurs, and an instance in which it does not occur, have every circumstance save one in common, that one occurring only in the former; the circumstance in which alone the two instances differ, is the effect, or cause, or a necessary part of the cause, of the phenomenon.*

In essence, the focus is on the comparison of outcomes between the units affected by the event of interest and those that remained unaffected, with a requirement for unaffected units to reproduce a counterfactual case, which would have happened, had there been no event or intervention occurring. Execution of such a strategy requires a careful choice of the comparison, or unaffected, unit in order to reduce bias and erroneous conclusions (Card and

Krueger (1994)) and since a suitable single comparison unit often does not exists, one would want to explore combinations of unaffected units to better reproduce characteristics of the affected (treated) unit.

The method that allows for such comparison unit selection constructs this unit as a weighted average of all potential comparison units with weights chosen in such a fashion that the difference in the characteristics between the synthesized unit (hence, the name of the method – the synthetic control) and the affected unit in the pre-intervention era is minimal (Abadie and Gardeazabal (2003)). In brief, the method specifies that the unobserved control (unaffected) outcome for the treated unit ($\hat{Y}(0)_{1,T}$) is imputed as $\hat{Y}_{1,T} = \sum_{i=2}^{N} w_i Y_{i,T}^{obs}$ where $i$ denotes 2 through N (untreated) units and scalar $w_i$ represents weight of $i^{th}$ untreated unit in the synthesized unit. To calculate the vector of weights $W = \{w_2 ... w_N\}$, the weighted least squares approach is used, so in the matrix form[1]:

$$\hat{W} = \underset{w}{\operatorname{argmin}} ||X_1 - X_0 W||v \tag{2.1}$$

subject to a number of constraints. Here $X_1$ represents the vector of pre-intervention characteristics for the treated unit which include the outcome variable $Y_1$ and a vector of observed covariates not affected by the intervention $Z_1$.[2]

The traditional version of the synthetic control method sis a constrained optimization problem, including such constraints as non-negative weights, no-intercept assumption, and restriction for weights to sum to one.[3]

[1] As per Abadie et al. (2007), $||X_1 - X_0W||v = \sqrt{(X_1 - X_0W)'V(X_1 - X_0W)}$

[2] Abadie et al. (2007), suggests a variety of specifications of the vector $X_1$. In particular, if the outcome variable in pre-intervention is observed over $T_0$ periods, let $K = (K_1, K_2, ...K_{T_0})'$ be a $(T_0 \times 1)$ vector of parameters used to construct a linear combination of the pre-intervention outcomes, so that $\overline{Y}_i^k = \sum_{s=1}^{T_0} k_s Y_{is}$ where $i = 1$ represents the treated unit and $i \in [2; N]$ represents the control units. Consider $M$ of such vectors: $K_1 ... K_M$. $X_1$ is then specified as $X_1 = (Z_1', \overline{Y}^{K_1}, ..., \overline{Y}^{K_M})$; $X_i$ where $i \in [2; N]$ representing control units, would be calculated in a similar way

[3] Non-negativity constraint specifies that $w_i \geq 0 \forall i \in (2 : N)$ and adding-up constraint specifies that $\sum(w_i) = 1$

I follow the Doudchenko and Imbens (2016) method and relax some assumptions initially imposed by Abadie and Gardeazabal (2003) and Abadie et al. (2007). First, I add an intercept $\mu$ to allow for possibility that the outcome for the treated unit is systematically larger by a constant amount; the need for this assumption is based on previous research that shows significant gap between trading book ratios and return volatility for treated and control group. Secondly, I do not restrict the weights to be nonnegative: banks profit is a result of its strategy that could be positively or negatively related to strategies of other banks that belong to a control group. The resulting model is an application of the elastic net regression as proposed by Zou and Hastie (2005).

The method takes a semi-parametric approach to pre-treatment outcomes by assigning an optimal weight to each control observation. A vector of weights, $w$, is obtained in a minimization problem:

$$(\hat{\mu}(\lambda,\alpha),\hat{w}(\lambda,\alpha)) = \underset{\mu,w}{\operatorname{argmin}} Q((\mu,w)|X^{obs}_{t,pre},X^{obs}_{c,pre};(\lambda,\alpha)), \tag{2.2}$$

$$Q((\mu,w)|X^{obs}_{t,pre},X^{obs}_{c,pre}) = ||X^{obs}_{t,pre} - \mu - w'X^{obs}_{c,pre}||^2_2 + \lambda(\frac{1-\alpha}{2}||w||^2_2 + \alpha||w||_1), \tag{2.3}$$

where $X^{obs}_{t,pre}, X^{obs}_{c,pre}$ are pre-intervention values[4] for treated and control. I select tuning parameters $(\lambda,\alpha)$ using the cross-validation method suggested by Breiman et al. (1984) to overcome over-fitting.

Such implementation design for an optimal vector of weights, $w^*$ should imply that the following holds approximately:

$$\sum_{i=2}^{N} w^*_i Y^{obs}_{i,pre} = Y^{obs}_{1,pre} \tag{2.4}$$

and

$$\sum_{j=2}^{N} w^*_i Z^{obs}_{i,pre} = Z^{obs}_{1,pre} \tag{2.5}$$

[4]in this case, $X_i = (Z'_i, Y_i)$

(where control index, $i \in [2:N]$) to ensure unbiasednes of the estimator.

Since large sample asymptotic inference is not possible for the synthetic control, the use of permutation methods is suggested (Abadie et al. (2015)). This requires estimation of a placebo treatment effect for each unit in a donor pool using synthetic control method, and calculating an empirical p-value for the effect estimated on the treatment unit. For this, define

$$\hat{\delta}_{i,\tau} = Y^{obs}_{i,\tau} - \hat{Y}^{obs}_{i,\tau}, \forall \tau \in [1:T] \tag{2.6}$$

as the estimated gap due to the enactment of the Rule. Here, $T$ is the last available month of the considered (pre- and post-) time period. Assuming that $T_0$ is the last month of pre-treatment period,

$$RMSPE_i = \frac{\sum_{t=T_0+1}^{T} (\hat{\delta}_{i,\tau})^2/(T-T_0)}{\sum_{t=1}^{T_0} (\hat{\delta}_{i,\tau})^2/(T_0)} \forall i \in [1:N] \tag{2.7}$$

The P-value is then calculated as

$$p = \frac{\sum_{i=2}^{N} I(RMSPE_i >= RMSPE_1)}{N}, \tag{2.8}$$

where $I$ is an indicator function given the condition in the parentheses.

## 2.4 Identification

The Volcker Rule's explicit aim is to reduce risk-taking by banks by limiting their prop trading and investment in hedge funds and private equity. The limitation to hold "covered funds" allows for a few exemptions, as discussed earlier, one of which stipulates that such investment must be "immaterial" to the bank and should constitute no more than 3% of Tier 1 Capital of the banking entity. The identification of affected banks based on this indicator was used in Keppo and Korte (2016) and Chung et al. (2019). Both studies, however, note that a 3% is rather arbitrary number for two reasons. Firstly, the Volcker Rule is still not fully implemented and is not yet fully binding for banks — some of them may have responded to a new regulation and made a public announcement about it, some are still in the process of compliance.

Secondly, the Rule specifies that the entities that could rely on exception from the definition of the investment company under the Investment Company Act could also be excluded from the definition of the covered fund (real estate funds, for example). This could enlarge the trading book (as defined in the aforementioned studies) without having a bank violate any of the Volcker stipulations. In this case, the noted ratio would be higher than the specified 3% yet the bank would not be affected by the law at all. Thus, application of such strategy would result in leaving a few "apples" in a basket of "lemons."

In addition, the considered identification strategy would not take into account the proprietary trading limitation, which forces banks to stop engaging as a principal for the trading account. It is true that some banks ceased all proprietary activity during or before 2010. However, some larger banks note in their financial reports that the portion of it remained after the Volcker enactment date and was expected to be exited by the Final Rules deadline which was in April, 2014.

To properly capture the legislation's effect on the banks' practices, one would need to find a reliable measure that could identify the line between the treatment (affected) and control (unaffected) pool. For instance, public announcement about compliance with the Volcker Rule made by a certain bank would not only indicate bank's assignment to treatment but potentially give some information about the start of the treatment. Publicly traded firms (banks) disclose their information on ongoing basis in compliance with the federal security laws. Thus, one can use text data from the annual 10K reports to extract firms' opinion about the effect of the Volcker Rule on them as well as progress they made towards meeting the Rule's requirements. This draws a distinct line between the banks that are subject to the Volcker Rule and those who are not. Accordingly, in this study, I follow the latter strategy.

## 2.5 Data

According to U.S. regulations, the banks that are in a status of a bank holding company are regulated by the Federal Reserve and required to file quarterly financial reporting on a consolidated level (FR Y-9C/LP/SP, available from the Chicago Fed). There is a limited number of banks that have long enough histories of quarterly financial reports. For instance, such banks as Goldman Sachs and Morgan Stanley switched to be the traditional bank holding companies only a year or two before the enactment of the Volcker Rule. There is a trade-off between the length of pre-intervention era chosen for the study and the number of banks to be included in it; however, both dimensions need to be maximized. A short pre-intervention period poses the small-training-set problem: the smaller size of training set increases the part of squared error of the model that is due to the variance term as defined by Geman et al. (1992); from the other side, a small pool of control units would negatively affect the fit of the model.

The time window considered in this study extends from 2001 Q1 to 2017Q3 and covers 120 banks. The data are obtained from the Bank Regulatory database of WRDS.

The variables of interest are the banks' non-interest revenues generated by trading and investment activities and defined according to the *Consolidated Income Statement - Report of Income for Holding Companies*. In particular, these are trading assets, securities available for sale, revenue from investment banking advisory, brokerage, and underwriting fees and commissions. The values are calculated as a ratios; see Table 1 for precise descriptions.

In addition, I consider gross notional amounts of derivative contracts held for trading as well as those held for purposes other than trading to proxy for the trading and hedge inventory levels. The changes in these levels could potentially be linked to the changes in hedging and market making activity with respect to this particular security class.

In the assessment of the riskiness of banks' activities, I follow previous studies. As a measure of a bank's risk, I consider a Z-score, calculated as the sum of Equity/TA and ROA divided by the standard deviation of ROA, estimated over rolling windows of two years (King, 2013) The set of covariates includes ratios covering profitability, efficiency, and asset quality. In particular, the controlling variables are RoE, Cost-to-Income ratio, and Loan Loss Provision to Gross Loans. The choice of the explanatory variables is limited due to the methodology requirements: covariates must be not affected by the legislation.

The flow variables in the BHC regulatory filings are reported on a year-to-date basis. Quarterly flow series are obtained by "quarterizing" the data: the variable at time (t-1) for a given year is subtracted from observations in quarter 2, 3, and 4 of the same year. I use a cubic spline imputation technique for the quarterly accounting data to derive monthly time series which are then merged with the 10-K reports text data using CIK-PERMCO links obtained from the WRDS database as well as PERMCO-RSSD links file provided by the Federal Reserve Bank of New York. The time period is 2001Q3 to 2017Q3 (the latest available for most banks), allowing for at least 88[5] observations before the treatment took place.

I extract annual 10K reports for 2010-2016 fiscal years from EDGAR (SEC) using Master Index directories. The public announcements with regard to the Volcker Rule compliance progress and firm's opinion about the Rule's impact on it are then located in the reports using text analysis. I scrape the html version of the 10-K filing for each bank and extract paragraphs with the Volcker Rule mentioned in them; I then choose only sentences with the bank's feedback about the legislation that is contained within this paragraph. The procedure

[5] Z-score calculation includes estimation of $\sigma$ of ROA which requires 2 years of past RoA data at each point; for the rest of variables, the pre-Volcker period constitutes 112 months

is repeated for 2011-2017 filing years that represent 2010-2016 fiscal years. The output is manually classified into the treated and not-treated pools. Those BHCs that clearly state that they are not a subject to the Rule constitute the control group; the BHCs in the treatment group note in their reports about changes to be done in compliance with the Volcker Rule and any material effect it might have on the business, financial condition, and results of operations. The sample size is 120 publicly traded banks out of which 47 noted to be affected by the Volcker Rule, and the other 73 reported that they neither held any covered funds, nor they engage in the trading activity prohibited by the Volcker Rule.

The smaller size treated banks are less likely to be involved in the prop trading activity in the pre-Volcker era. The primary change in the operation of such banks in compliance with the regulation would be a sell-off of the prohibited covered funds, as they mention in their 10-K reports. This is consistent with the recent calculations presented by the Federal Agencies (OCC et al. (2018)), showing that roughly 98 percent of the trading assets and liabilities in the U.S. banking system are currently held by firms with trading assets and liabilities of 1 billion or more. Thus, it is reasonable to distinguish between smaller and larger banks when running aggregate analysis. For that purpose, I construct representative big and small bank. I define a big bank as one that has over $50 billion in total assets. The representative big bank is then constructed as the total-asset-weighted average of all treated banks populating the big banks group; the leftover treated banks are used in the calculation of the representative small bank.

Figure 1 depicts simple averages for the covariates (RoE, Cost-to-Income ratio, and Loan Loss Provision to Gross Loans) across banks in the treated and control groups. None of the variables shows long-lasting differences between the groups. Summaries of the pre- and post-treatment averages for each of the variables are presented in Table 2 for aggregate small, big, and control bank: no significant change across the representative banks is noted.

Table 3 summarizes the variables of interest. Trading Assets constitute slightly more than

1% of Total Assets for the small representative bank which see a marginally small increase in this accounting item post-Volcker; the representative big bank, on the other hand, allocates 14% of its total assets to the trading activity in the pre-Volcker period. This number drops by 2.7% in the post-intervention. At the same time, the big bank almost doubles its total notional amounts of derivatives held for the purpose other then trading, the derivatives held for trading increase significantly too for both small and big banks, however, in comparison with the control bank, such change is not dramatic. The increase in the Z-score for big banks is much more pronounced for the big bank compared to both the small bank and the control bank, potentially pointing to the negative impact of the Volcker Rule on the riskiness of those "too big to fail."

## 2.6 Estimation and results

For the estimation purposes all variables are normalized using estimates of means and standard deviations that are calculated based on the pre-treatment (train) set for each of treatment and control banks. The estimation process includes an automatic choice of tuning parameters, ($\alpha$, $\lambda$) via the cros-validation procedure; I follow the suggestion of Doudchenko and Imbens (2016) to consider all positive values of $\lambda$ on a grid of $\alpha \in \{0.1, ..., 0.9\}$ The cross-validation procedure is then carried out as following: for each $\alpha$, such $\lambda$ is calculated that minimizes mean squared error, denoted as $CV(\lambda)$:

$$CV_\alpha(\lambda) = \frac{1}{T - T_0} \sum_{k=1}^{5} \sum_{i=T_0+1}^{T} (\hat{\delta}_{k,t}^{(-k)})^2 \tag{2.9}$$

where $k$ represents the folds[6]: a vector of treated values, $Y_{t,pre}^{obs}$, is divided into five folds and for each of these folds the model is fitted using the other four (denoted in the Equation (9)

[6]the number of folds is traditionally set to 5.

by superscript $(-k)$ meaning non-$k^{th}$ folds) as a training set. Finally, $\alpha$ is chosen for which $CV_\alpha(\lambda)$ is minimal. Essentially, the optimal pair of $(\alpha, \lambda)$ that minimizes $CV()$ on a given grid of $\alpha$ is chosen.

To obtain a more parsimonious model, I follow Breiman (1984) method that suggests to choose $\lambda$ for which

$$CV_{\hat{\lambda}} = \min_{\lambda} CV(\lambda) + s(\min_{\lambda} CV(\lambda)) \tag{2.10}$$

where $s()$ is a standard error operator (standard error is estimated based on five values of mean squared error obtained from fitting the model for each fold). In essence, I choose the simplest model — that yields the smallest $CV$ — which is no more than one standard error worse than the best model.

Using elastic net regression, I construct at most 47 synthetic treated units[7] for each variable of interest. In addition, in order to summarize the results, I repeat the estimation of the synthetic control for a representative big and small bank, constructed as noted earlier.

The estimation procedure is carried over sets of controls populated by 18 to 73 banks for the representative big and small treated units as well as sets of 25 to 47 separate affected banks. The seminal synthetic control study (Abadie and Gardeazabal (2003)) imposes restriction of weights summing to one on the optimization process, as an assurance of stability of the model. The elastic net approach allows to relax this constraint by imposing a combination of LASSO and Ridge regression constraints. However, in this application, the sum of weights for each considered case does not fall far from 1. The weights are presented in Tables 8 – 11.

In assessment of the Rule's impact on banks, I define two post-intervention periods: pre-Final Rule and post-Final Rule. The Final Rule was adopted in April 2014, coming into effect

[7] Some of the treated banks do not show any activity with respect to a particular variable, especially in case when it is a smaller bank (for instance some smaller banks do not trade but do report presence of covered funds that need to be liquidated); the numbers of observed banks in each category (small, big, or control) for each variable are presented in Table 3.

in July 2015, and included a set of revisions. In particular, it reduced the burden on smaller and less-complex institutions by shrinking their compliance and reporting requirements. In addition, it permitted banks to retain certain CDOs, and gave an extension in conformance with the covered funds clause.

Krawiec and Liu (2015) note that the effort to influence the Rule at the agency level began early in the legislation process. After the enactment in 2010 the Financial Stability Oversight Council (FSOC) committee was formed to conduct a study and make recommendations on effective Volcker Rule implementation. Prior to the study, the FSOC solicited public input, and roughly 18450 comments were received. The comments addressed all major provisions of the proposed Rule; however, major attention was paid to the market-making exemption. For instance, the commenters worried that the Volcker Rule, in its original form (the Proposal), would limit banks' ability to engage in market-making, which would negatively affect market liquidity, price discovery, and capital formation.

The Final Rule took into account the comments and made substantial revisions to treatment of the market making and hedging activities. In particular, under the Final Regulation, a trading desk may engage in the market-making activity without trade-by-trade analysis required under the Proposal, but instead a banking entity must monitor its financial exposure (aggregate risks of the financial instruments which the desk trades) and market-making inventory; the Proposal had a provision which would require market-making to generate revenues primarily from fees, commissions, bid-ask spread, or any other source except the appreciation in the value of covered financial position, which by many deemed to be quite restrictive. This provision was not included into the Final Rule.

While relaxing regulation of market-making, the Final Rule imposed additional compliance obligations on the hedging activity: under the Final regulation, for each position to qualify as a hedge, some type of analysis should demonstrate that this position can mitigate specific identifiable risk. In addition, hedging could no longer be used to reduce risks associated with the general market movements, broad economic conditions, or other macro-based risk.

The mentioned changes in the rule-making process required this study to separate between pre- (June 2010 - March 2014) and post-Final Rule (April 2014 and onward) period.

Consistent with previous studies (Keppo and Korte (2016), Chung et al. (2019)), I find reduction in trading asset compared to constructed counterfactual for the representative big bank. Figure 2 (panel (a)) presents the true and counterfactual values for the considered variable; it shows the widening gap between the two, suggesting a lingering effect of the legislation on the banks' assets structure. The result, however, is not statistically significant with respect to the standard confidence levels; no significant change was captured for the representative smaller bank (Figure 3).

Tables 14 and 15 show the estimated gaps between synthetic and observed values for each of treated banks. The biggest banks experience slight although not significant shrinkage of their trading assets. Few banks (Wells Fargo, U.S. Bancorp), however, experience a statistically significant increase in this accounting item.

For the treated representative bank whose total assets are bigger than 50 billion, revenue generated by the IB advisory, brokerage, and underwriting increased slightly although such change was not statistically significant (Figure 2, panel (b); Table 12). In the bank-by-bank analysis (Tables 16 and 17), in the post-Final Rule period, nine treated banks show significant increase in this item, including such banks as Bank of America and Wells Fargo. This finding suggests that, relative to the other lines of business, for some banks, this revenue-generating practice potentially gained more importance in the post-Volcker era.

As Figure 2, panel(c) and Table 20 show, some of the biggest banks (Bank of America, JPMorgan) have expanded their derivative hedging activities – the total gross notional amounts of derivatives held for the purpose other than trading have increased drastically over the pre-Final Rule period. As noted earlier, prop trading ban has a vast list of permitted exceptions, including hedging and market-making, and differentiation between prohibited and permitted activity often is hard to make. A broken hedging strategy could resemble proprietary trading. Likewise, such short-term increase in this financial report item could possibly signify an attempt of presenting banned activity as a legitimate hedging strategy - the sharp increase only happens in pre-Final Rule period and becomes insignificant after the final, more stringent (with respect to hedging), regulation is released. Alternatively, a potential explanation of such increase could be a growing need of managing inventory risk for bank as a market-maker.

The Volcker Rule aimed for a reduction of riskiness of the banking actions and existing literature documents it being ineffective with that respect: Chung et al. (2019) find that the Rule has raised default probability of the affected banks. Figure 2 (panel (e)) shows small positive gap between true and synthetic Z-score for a representative big bank; however, statistically speaking, this gap is not different from zero. As in case of trading assets, the distance between the two widens with time: the drop in riskiness of banking activity is statistically significant over the last quarter of the post-intervention window at 90% confidence level. In the bank-by-bank analysis, four out of 15 big banks, including Citigroup and JPM Chase, experience significant increase in the Z-score in the post-Final Rule era. The effect of the Rule on the small representative bank is virtually zero (Figure 3, panel(e), Table 23).

In light of recent developments around the Rule, I also consider changes in available-for-sale securities. In May 2018, the Board issued proposed revisions to the regulation which should potentially simplify the Rule and reduce compliance costs without negatively affecting safety and soundness of the banking entities. One of the proposed changes included revision of the term "trading account" by replacing short-term intent-based prong with the accounting-based prong, which in general would cover derivatives, trading securities, and available-for-sale securities. The proposal was met with negative feedback from the banks, as the available-for-sale category of securities is quite broad and includes investment grade corporate bonds and asset backed securities.

If the accounting-based prong coincides with the previously adopted prong, given that the Volcker Rule had at least small effect on the banks, some changes in the available-for-sale securities category should be captured. Figure 2 (segment (f)) shows small increase (although not lasting) for the representative big bank and decrease for the representative small bank in both cases with overwhelmingly high p-values, suggesting a need for finer classification in determining potentially affected accounting items.

## 2.7 Conclusion

This study analyzes the impact the Volcker Rule had on the U.S. bank holding companies. A new identification strategy was designed, based on the text portion of the banks' annual financial reporting. The assumptions behind this study are built on previous findings (as well as the finding in this study) that suggested that the Rule had no impact on the banks' profits; in a sense profitability is an exogenous measure, a level that the bank aims to meet so its stock performance is not penalized. Consistent with previous studies, I find change in assets structure in the banks in compliance with the regulation. The notional value of the derivatives held for the purpose other than trading increases drastically over the pre-Final Rule period. Such activity coincides with the heated interest from the banks and public to relax the initial Rule's constrains on the market-making. The gap between the true and counterfactual derivatives diminishes significantly in the post-Final rule period. Finally, previous studies find that the Rule was not effective with respect to its risk-decreasing aim. My result is consistent with the finding only over the portion of the considered post-intervention window, suggesting that the regulation did bring some positive changes into the current financial system.

Table 2.1: Variable Definitions

| Variable | Bank regulatory database item |
|---|---|
| Trading assets | BHCK3545/BHCK2170 |
| Investmet banking advisory, brokerage, and underwriting fees and commissons | (BHCKC386 + BHCKC387 + BHCKC886 + BHCKC887 + BHCKC888)/(BHCK4074+BHCK4079) |
| Total gross notional amount of derivative contracts held for trading | BHCKA126 + BHCKA127 + BHCK8723 + BHCK8724 |
| Total gross notional amount of derivative contracts held for purpose other than trading | BHCK8725 + BHCK8726 + BHCK8727 + BHCK8728 |
| Z-score | ((BHCK3519/BHCK2170)+RoA)/sRoA |
| Available for sale securities | BHCK1773/BHCK2170 |
| Efficiency (Cost-Income ratio) | BHCK4093/(BHCK4107 + BHCK4079) |
| Profitability (RoE) | BHCK4340 / BHCK2170 |

Table 2.2: Summary statistics. Covariates

| Variable | Pre-Volcker | Post-Volcker | Change | N of banks |
|---|---|---|---|---|
| | Small Bank | | | |
| Profitability | 0.002 | 0.002 | 0.000 | 32 |
| Efficiency | 0.683 | 0.676 | -0.007 | 32 |
| Asset Quality | 0.001 | 0.000 | -0.001 | 32 |
| | Big Bank | | | |
| Profitability | 0.002 | 0.002 | 0.000 | 15 |
| Efficiency | 0.643 | 0.691 | 0.048 | 15 |
| Asset Quality | 0.001 | 0.000 | -0.001 | 15 |
| | Control Bank | | | |
| Profitability | 0.002 | 0.002 | 0.000 | 73 |
| Efficiency | 0.640 | 0.665 | 0.025 | 73 |
| Asset Quality | 0.001 | 0.000 | -0.001 | 73 |

*Big banks are defined as such that have total assets of 50B and more (based on pre-intervention); small banks are defined as all other banks populating the treated group. All values are weighted by the total assets

Table 2.3: Summary statistics. Variables of Interest

| Variable | Pre-Volcker | Post-Volcker | Change | N of banks |
|---|---|---|---|---|
| | Small Bank | | | |
| Trading assets (as % of total assets) | 0.012 | 0.013 | 0.0013 | 16 |
| Investment banking advisory, brokerage, and underwriting fees and commissions (as % of interest and noninterest income) | 0.023 | 0.0308 | 0.007 | 30 |
| Total GNA of derivative contracts held for trading | 4261852 | 8277239 | 4015387 | 10 |
| Total GNA of derivative contracts held for purpose other than trading | 3103719 | 1055982 | -2047737 | 24 |
| Z-score | 272 | 330.7 | 58.7 | 30 |
| Available for sale securities (as % of total assets) | 0.192 | 0.195 | 0.0032 | 32 |
| | Big Bank | | | |
| Trading assets (as % of total assets) | 0.140 | 0.113 | -0.02695 | 15 |
| Investment banking advisory, brokerage, and underwriting fees and commissions (as % of interest and noninterest income) | 0.0761 | 0.131 | 0.05473 | 15 |
| Total GNA of derivative contracts held for trading | 2.40E+10 | 3.90E+10 | 1.50E+10 | 15 |
| Total GNA of derivative contracts held for purpose other than trading | 3.20E+08 | 5.80E+08 | 2.60E+08 | 15 |
| Z-score | 183.59 | 269.92 | 86.33 | 15 |
| Available for sale securities (as % of total assets) | 0.130 | 0.147 | 0.017 | 15 |
| | Control Bank | | | |
| Trading assets (as % of total assets) | 0.007 | 0.0053 | -0.0021 | 18 |
| Investment banking advisory, brokerage, and underwriting fees and commissions (as % of interest and noninterest income) | 0.0621 | 0.07 | 0.0079 | 72 |
| Total GNA of derivative contracts held for trading | 1.10E+07 | 2.90E+07 | 1.80E+07 | 24 |
| Total GNA of derivative contracts held for purpose other than trading | 9.00E+06 | 1.60E+07 | 7.20E+06 | 50 |
| Z-score | 298.3 | 386.07 | 69.746 | 69 |
| Available for sale securities (as % of total assets) | 0.182 | 0.174 | -0.0076 | 73 |

*Big banks are defined as such that have total assets of 50B and more (based on pre-intervention); small banks are defined as all other banks populating the treated group. All values are weighted by the total assets. The dollar amounts are reported in thousands of dollars

Table 2.4: Summary statistics. Covariates. Control banks

| | Profitability | | | Efficiency | | | Asset quality | | |
|---|---|---|---|---|---|---|---|---|---|
| Banks | Pre- | Post- | Change | Pre- | Post- | Change | Pre- | Post- | Change |
| 1St Source Corporation | 0.0019 | 0.0028 | 0.0009 | 0.7085 | 0.6424 | -0.0661 | 0.0003 | 0.0000 | -0.0003 |
| American National Bankshares Inc. | 0.0034 | 0.0027 | -0.0008 | 0.5563 | 0.6343 | 0.0780 | 0.0002 | 0.0001 | -0.0001 |
| Associated Banc-Corp | 0.0028 | 0.0017 | -0.0010 | 0.5273 | 0.7009 | 0.1736 | 0.0012 | -0.0001 | -0.0013 |
| Bancfirst Corporation | 0.0030 | 0.0025 | -0.0005 | 0.6491 | 0.6544 | 0.0054 | 0.0001 | 0.0001 | 0.0000 |
| Bank Of Hawaii Corporation | 0.0038 | 0.0029 | -0.0009 | 0.5950 | 0.5928 | -0.0022 | 0.0002 | 0.0000 | -0.0003 |
| Bar Harbor Bankshares | 0.0022 | 0.0024 | 0.0002 | 0.6642 | 0.6020 | -0.0622 | 0.0002 | 0.0001 | 0.0000 |
| Bb&T Corporation | 0.0032 | 0.0025 | -0.0008 | 0.5646 | 0.6338 | 0.0692 | 0.0007 | 0.0003 | -0.0004 |
| C&F Financial Corporation | 0.0039 | 0.0030 | -0.0009 | 0.6740 | 0.6505 | -0.0235 | 0.0010 | 0.0014 | 0.0004 |
| Capital City Bank Group, Inc. | 0.0024 | 0.0007 | -0.0018 | 0.7092 | 0.8710 | 0.1617 | 0.0005 | 0.0002 | -0.0003 |
| Center Bancorp, Inc. | 0.0016 | 0.0025 | 0.0009 | 0.7290 | 0.5232 | -0.2058 | 0.0002 | 0.0005 | 0.0002 |
| Central Pacific Financial Corp. | -0.0007 | 0.0026 | 0.0033 | 0.7623 | 0.7670 | 0.0047 | 0.0014 | -0.0009 | -0.0023 |
| Century Bancorp, Inc. | 0.0015 | 0.0015 | 0.0000 | 0.7170 | 0.7098 | -0.0072 | 0.0003 | 0.0001 | -0.0002 |
| City Holding Company | 0.0039 | 0.0036 | -0.0002 | 0.5270 | 0.5477 | 0.0206 | 0.0001 | 0.0002 | 0.0001 |
| Columbia Banking System, Inc. | 0.0021 | 0.0026 | 0.0005 | 0.6611 | 0.6803 | 0.0191 | 0.0007 | 0.0000 | -0.0007 |
| Commerce Bancshares, Inc. | 0.0035 | 0.0030 | -0.0005 | 0.6078 | 0.6151 | 0.0073 | 0.0005 | 0.0002 | -0.0003 |
| Cullen/Frost Bankers, Inc. | 0.0035 | 0.0028 | -0.0008 | 0.6125 | 0.6480 | 0.0355 | 0.0002 | 0.0002 | 0.0001 |
| Cvb Financial Corp. | 0.0034 | 0.0033 | -0.0001 | 0.5177 | 0.5037 | -0.0139 | 0.0004 | -0.0001 | -0.0005 |
| Enterprise Bancorp, Inc. | 0.0022 | 0.0019 | -0.0002 | 0.7102 | 0.7108 | 0.0006 | 0.0003 | 0.0002 | -0.0001 |
| Enterprise Financial Services Corp | 0.0008 | 0.0024 | 0.0017 | 0.7642 | 0.5971 | -0.1671 | 0.0007 | 0.0001 | -0.0006 |
| Farmers Capital Bank Corporation | 0.0016 | 0.0017 | 0.0002 | 0.7798 | 0.7684 | -0.0115 | 0.0006 | 0.0000 | -0.0006 |
| Farmers National Banc Corp. | 0.0022 | 0.0022 | -0.0001 | 0.6433 | 0.7240 | 0.0807 | 0.0004 | 0.0002 | -0.0002 |
| Financial Institutions, Inc. | 0.0015 | 0.0023 | 0.0008 | 0.6503 | 0.6243 | -0.0260 | 0.0006 | 0.0004 | -0.0002 |
| First Bancorp | 0.0013 | 0.0002 | -0.0011 | 0.5342 | 0.7099 | 0.1757 | 0.0012 | 0.0009 | -0.0003 |
| First Busey Corporation | 0.0001 | 0.0022 | 0.0021 | 0.7279 | 0.6553 | -0.0726 | 0.0014 | 0.0002 | -0.0012 |
| First Citizens Banc Corp | 0.0008 | 0.0017 | 0.0009 | 0.7989 | 0.7217 | -0.0772 | 0.0006 | 0.0001 | -0.0005 |
| First Citizens Bancshares, Inc. | 0.0018 | 0.0018 | 0.0000 | 0.7265 | 0.7172 | -0.0093 | 0.0003 | 0.0005 | 0.0002 |
| First Commonwealth Financial Corporation | 0.0018 | 0.0018 | 0.0000 | 0.6573 | 0.6718 | 0.0145 | 0.0005 | 0.0002 | -0.0002 |
| First Community Bancshares, Inc. | 0.0024 | 0.0025 | 0.0001 | 0.6827 | 0.6486 | -0.0341 | 0.0004 | 0.0002 | -0.0003 |
| First Financial Corporation | 0.0029 | 0.0030 | 0.0001 | 0.5962 | 0.6248 | 0.0287 | 0.0004 | 0.0002 | -0.0002 |
| First Merchants Corporation | 0.0018 | 0.0024 | 0.0006 | 0.6519 | 0.6613 | 0.0094 | 0.0006 | 0.0001 | -0.0005 |
| First Mid Ill Bancshares Inc | 0.0026 | 0.0022 | -0.0004 | 0.6455 | 0.6472 | 0.0017 | 0.0002 | 0.0001 | -0.0001 |
| First Midwest Bancorp, Inc. | 0.0027 | 0.0014 | -0.0013 | 0.5572 | 0.6926 | 0.1354 | 0.0010 | 0.0006 | -0.0004 |
| Fulton Financial Corporation | 0.0029 | 0.0023 | -0.0006 | 0.5924 | 0.6433 | 0.0509 | 0.0004 | 0.0002 | -0.0002 |
| German American Bancorp, Inc. | 0.0024 | 0.0031 | 0.0007 | 0.6870 | 0.6052 | -0.0818 | 0.0002 | 0.0001 | -0.0001 |
| Hanmi Financial Corporation | -0.0002 | 0.0034 | 0.0037 | 0.6391 | 0.6047 | -0.0344 | 0.0017 | -0.0006 | -0.0023 |
| Heartland Financial Usa, Inc. | 0.0018 | 0.0021 | 0.0003 | 0.7091 | 0.7355 | 0.0265 | 0.0006 | 0.0002 | -0.0004 |

Table 2.5: Summary statistics. Covariates. Control banks. Continued

| | Profitability | | | Efficiency | | | Asset quality | | |
|---|---|---|---|---|---|---|---|---|---|
| Banks | Pre- | Post- | Change | Pre- | Post- | Change | Pre- | Post- | Change |
| Heritage Commerce Corp | 0.0009 | 0.0017 | 0.0008 | 0.7673 | 0.7276 | -0.0397 | 0.0007 | -0.0005 | -0.0011 |
| Home Bancshares, Inc. | 0.0019 | 0.0037 | 0.0018 | 0.6010 | 0.4655 | -0.1354 | 0.0007 | 0.0011 | 0.0004 |
| Huntington Bancshares Incorporated | 0.0003 | 0.0025 | 0.0022 | 0.7613 | 0.6615 | -0.0998 | 0.0019 | 0.0000 | -0.0018 |
| Iberiabank Corporation | 0.0028 | 0.0017 | -0.0011 | 0.6124 | 0.7472 | 0.1348 | 0.0004 | 0.0001 | -0.0002 |
| Independent Bank Corp. | 0.0025 | 0.0024 | -0.0001 | 0.6434 | 0.6619 | 0.0184 | 0.0003 | 0.0001 | -0.0002 |
| Independent Bank Corporation | 0.0006 | 0.0024 | 0.0017 | 0.7256 | 0.8129 | 0.0873 | 0.0010 | -0.0001 | -0.0011 |
| International Bancshares Corporation | 0.0033 | 0.0028 | -0.0006 | 0.5536 | 0.5863 | 0.0327 | 0.0002 | 0.0003 | 0.0001 |
| Lakeland Bancorp, Inc. | 0.0019 | 0.0021 | 0.0002 | 0.6324 | 0.6066 | -0.0258 | 0.0005 | 0.0002 | -0.0003 |
| Lakeland Financial Corporation | 0.0026 | 0.0031 | 0.0005 | 0.5815 | 0.4915 | -0.0900 | 0.0004 | 0.0001 | -0.0003 |
| Midsouth Bancorp, Inc. | 0.0026 | 0.0014 | -0.0011 | 0.7299 | 0.7571 | 0.0272 | 0.0004 | 0.0006 | 0.0002 |
| Midwestone Financial Group, Inc. | 0.0017 | 0.0023 | 0.0005 | 0.7047 | 0.6543 | -0.0504 | 0.0002 | 0.0003 | 0.0001 |
| Nara Bancorp, Inc. | 0.0028 | 0.0029 | 0.0001 | 0.5515 | 0.4891 | -0.0624 | 0.0013 | -0.0003 | -0.0016 |
| Nbt Bancorp Inc. | 0.0026 | 0.0024 | -0.0002 | 0.5894 | 0.6474 | 0.0580 | 0.0007 | 0.0003 | -0.0004 |
| Park National Corporation | 0.0035 | 0.0027 | -0.0007 | 0.5412 | 0.6055 | 0.0644 | 0.0007 | 0.0003 | -0.0004 |
| Peoples Bancorp Inc. | 0.0023 | 0.0019 | -0.0004 | 0.6092 | 0.7380 | 0.1288 | 0.0008 | 0.0001 | -0.0007 |
| Peoples Utah Bancorp | 0.0038 | 0.0028 | -0.0010 | 0.5269 | 0.6195 | 0.0926 | 0.0009 | 0.0001 | -0.0008 |
| Popular, Inc. | 0.0004 | 0.0021 | 0.0018 | 0.6497 | 0.8373 | 0.1876 | 0.0016 | 0.0010 | -0.0006 |
| Prosperity Bancshares, Inc. | 0.0033 | 0.0034 | 0.0000 | 0.4805 | 0.4232 | -0.0573 | 0.0002 | 0.0001 | -0.0001 |
| Qcr Holdings, Inc. | 0.0013 | 0.0017 | 0.0004 | 0.7418 | 0.7062 | -0.0356 | 0.0005 | 0.0003 | -0.0002 |
| Regions Financial Corporation | 0.0010 | 0.0017 | 0.0007 | 0.7595 | 0.6765 | -0.0830 | 0.0011 | 0.0002 | -0.0009 |
| Renasant Corporation | 0.0025 | 0.0022 | -0.0003 | 0.6718 | 0.6937 | 0.0219 | 0.0004 | 0.0002 | -0.0003 |
| Republic Bancorp, Inc. | 0.0031 | 0.0035 | 0.0004 | 0.6047 | 0.6562 | 0.0515 | 0.0003 | 0.0000 | -0.0002 |
| Royal Bancshares Of Pennsylvania, Inc. | 0.0013 | 0.0000 | -0.0014 | 0.5557 | 0.9477 | 0.3920 | 0.0010 | 0.0010 | 0.0000 |
| S & T Bancorp, Inc. | 0.0040 | 0.0027 | -0.0013 | 0.4673 | 0.5864 | 0.1191 | 0.0003 | 0.0002 | -0.0001 |
| S. Y. Bancorp, Inc. | 0.0038 | 0.0034 | -0.0004 | 0.5790 | 0.5865 | 0.0076 | 0.0003 | 0.0002 | -0.0001 |
| Sandy Spring Bancorp, Inc. | 0.0023 | 0.0025 | 0.0002 | 0.6669 | 0.6383 | -0.0287 | 0.0006 | 0.0000 | -0.0006 |
| Scbt Financial Corporation | 0.0027 | 0.0021 | -0.0006 | 0.6430 | 0.6971 | 0.0541 | 0.0005 | 0.0001 | -0.0005 |
| Shore Bancshares, Inc. | 0.0031 | 0.0002 | -0.0029 | 0.5732 | 0.7289 | 0.1557 | 0.0004 | 0.0006 | 0.0002 |
| Southwest Georgia Financial Corporation | 0.0023 | 0.0018 | -0.0005 | 0.7594 | 0.8036 | 0.0442 | 0.0003 | 0.0001 | -0.0002 |
| Summit Financial Group, Inc. | 0.0018 | 0.0016 | -0.0003 | 0.6373 | 0.6616 | 0.0243 | 0.0005 | 0.0001 | -0.0005 |
| Texas Capital Bancshares, Inc. | 0.0017 | 0.0025 | 0.0008 | 0.6777 | 0.5477 | -0.1301 | 0.0004 | 0.0001 | -0.0003 |
| Tompkins Financial Corporation | 0.0032 | 0.0025 | -0.0007 | 0.6118 | 0.6504 | 0.0386 | 0.0002 | 0.0001 | -0.0001 |
| Trustmark Corporation | 0.0034 | 0.0025 | -0.0008 | 0.5698 | 0.6959 | 0.1261 | 0.0003 | 0.0001 | -0.0002 |
| Union First Market Bankshares Corporation | 0.0027 | 0.0021 | -0.0006 | 0.6766 | 0.6988 | 0.0222 | 0.0003 | 0.0002 | -0.0001 |
| United Community Banks, Inc. | 0.0007 | -0.0004 | -0.0012 | 0.6820 | 0.9306 | 0.2486 | 0.0014 | 0.0000 | -0.0014 |
| Valley National Bancorp | 0.0033 | 0.0020 | -0.0013 | 0.5172 | 0.6648 | 0.1476 | 0.0002 | 0.0001 | -0.0001 |
| West Bancorporation, Inc. | 0.0032 | 0.0031 | -0.0001 | 0.4207 | 0.5181 | 0.0974 | 0.0003 | 0.0000 | -0.0003 |

Table 2.6: Summary statistics. Covariates. Small banks

| | Profitability | | | Efficiency | | | Asset quality | | |
|---|---|---|---|---|---|---|---|---|---|
| Banks | Pre- | Post- | Change | Pre- | Post- | Change | Pre- | Post- | Change |
| Auburn National Bancorporation, Inc. | 0.0024 | 0.0022 | -0.0002 | 0.5802 | 0.6319 | 0.0518 | 0.0004 | 0.0001 | -0.0003 |
| Banner Corporation | 0.0002 | 0.0017 | 0.0015 | 0.7737 | 0.7309 | -0.0428 | 0.0008 | 0.0001 | -0.0006 |
| Bok Financial Corporation | 0.0028 | 0.0027 | -0.0002 | 0.6069 | 0.6573 | 0.0504 | 0.0005 | -0.0001 | -0.0007 |
| Boston Private Financial Holdings, Inc. | 0.0005 | 0.0020 | 0.0016 | 0.8473 | 0.7506 | -0.0967 | 0.0004 | 0.0001 | -0.0004 |
| Camden National Corporation | 0.0031 | 0.0024 | -0.0006 | 0.5331 | 0.6149 | 0.0818 | 0.0002 | 0.0001 | -0.0001 |
| Cathay General Bancorp | 0.0030 | 0.0029 | -0.0001 | 0.4147 | 0.5082 | 0.0935 | 0.0009 | -0.0003 | -0.0012 |
| Chemung Financial Corp | 0.0024 | 0.0018 | -0.0005 | 0.7049 | 0.7671 | 0.0622 | 0.0003 | 0.0001 | -0.0001 |
| Colony Bankcorp, Inc. | 0.0012 | 0.0012 | 0.0000 | 0.6402 | 0.7439 | 0.1037 | 0.0011 | 0.0002 | -0.0010 |
| Community Bank System, Inc. | 0.0025 | 0.0029 | 0.0004 | 0.6804 | 0.6703 | -0.0101 | 0.0003 | 0.0002 | -0.0001 |
| F.N.B. Corporation | 0.0023 | 0.0023 | -0.0001 | 0.6494 | 0.6298 | -0.0196 | 0.0006 | 0.0002 | -0.0004 |
| First Financial Bancorp | 0.0035 | 0.0025 | -0.0010 | 0.6734 | 0.6171 | -0.0563 | 0.0007 | 0.0002 | -0.0005 |
| First Financial Bankshares, Inc. | 0.0043 | 0.0042 | -0.0001 | 0.5422 | 0.5265 | -0.0157 | 0.0003 | 0.0001 | -0.0002 |
| First Horizon National Corporation | 0.0020 | 0.0012 | -0.0008 | 0.7575 | 0.8428 | 0.0853 | 0.0009 | 0.0000 | -0.0009 |
| First United Corporation | 0.0017 | 0.0011 | -0.0006 | 0.6526 | 0.7842 | 0.1316 | 0.0005 | 0.0003 | -0.0002 |
| Heritage Financial Corporation | 0.0028 | 0.0023 | -0.0006 | 0.6098 | 0.6952 | 0.0855 | 0.0005 | 0.0002 | -0.0003 |
| Investors Bancorp, Mhc | 0.0006 | 0.0016 | 0.0010 | 0.6432 | 0.5161 | -0.1271 | 0.0002 | 0.0003 | 0.0001 |
| Mainsource Financial Group, Inc. | 0.0016 | 0.0024 | 0.0008 | 0.7079 | 0.7098 | 0.0019 | 0.0006 | 0.0000 | -0.0006 |
| Old National Bancorp | 0.0020 | 0.0024 | 0.0003 | 0.7237 | 0.7461 | 0.0223 | 0.0005 | 0.0000 | -0.0005 |
| Old Second Bancorp, Inc. | 0.0015 | 0.0010 | -0.0005 | 0.6871 | 0.8473 | 0.1602 | 0.0012 | -0.0005 | -0.0017 |
| Pacwest Bancorp | -0.0017 | 0.0029 | 0.0046 | 0.9023 | 0.5789 | -0.3234 | 0.0008 | -0.0002 | -0.0011 |
| Rockville Financial, Inc. | 0.0010 | 0.0016 | 0.0005 | 0.7003 | 0.7484 | 0.0481 | 0.0003 | 0.0002 | -0.0001 |
| Seacoast Banking Corporation Of Florida | -0.0006 | 0.0013 | 0.0018 | 0.7667 | 0.8645 | 0.0978 | 0.0015 | -0.0001 | -0.0017 |
| Southside Bancshares, Incorporated | 0.0027 | 0.0026 | -0.0001 | 0.6637 | 0.6551 | -0.0085 | 0.0005 | 0.0004 | -0.0001 |
| Sun Bancorp, Inc | 0.0005 | -0.0018 | -0.0023 | 0.8176 | 1.0257 | 0.2081 | 0.0005 | 0.0007 | 0.0002 |
| Svb Financial Group | 0.0033 | 0.0022 | -0.0011 | 0.6834 | 0.5489 | -0.1345 | 0.0004 | 0.0005 | 0.0001 |
| Synovus Financial Corp. | 0.0017 | 0.0018 | 0.0001 | 0.7477 | 0.6812 | -0.0665 | 0.0013 | 0.0003 | -0.0010 |
| Trico Bancshares | 0.0029 | 0.0023 | -0.0006 | 0.6222 | 0.6701 | 0.0479 | 0.0005 | 0.0001 | -0.0004 |
| Umb Financial Corporation | 0.0021 | 0.0020 | -0.0001 | 0.7963 | 0.7764 | -0.0199 | 0.0003 | 0.0002 | -0.0001 |
| United Bankshares, Inc. | 0.0033 | 0.0026 | -0.0007 | 0.5199 | 0.5405 | 0.0206 | 0.0002 | 0.0001 | -0.0001 |
| Washington Trust Bancorp, Inc. | 0.0024 | 0.0028 | 0.0005 | 0.6483 | 0.6206 | -0.0277 | 0.0001 | 0.0001 | 0.0000 |
| Wesbanco, Inc. | 0.0023 | 0.0024 | 0.0000 | 0.6333 | 0.6250 | -0.0083 | 0.0005 | 0.0002 | -0.0003 |
| Wintrust Financial Corporation | 0.0018 | 0.0018 | 0.0000 | 0.6591 | 0.6589 | -0.0002 | 0.0004 | 0.0002 | -0.0003 |

Table 2.7: Summary statistics. Covariates. Big banks

| | Profitability | | | Efficiency | | | Asset quality | | |
|---|---|---|---|---|---|---|---|---|---|
| Banks | Pre- | Post- | Change | Pre- | Post- | Change | Pre- | Post- | Change |
| Bank Of America Corporation | 0.0026 | 0.0010 | -0.0016 | 0.5496 | 0.8215 | 0.2719 | 0.0013 | 0.0001 | -0.0011 |
| Citigroup Inc. | 0.0020 | 0.0016 | -0.0004 | 0.7628 | 0.6618 | -0.1010 | 0.0019 | 0.0006 | -0.0012 |
| Citizens Financial Group Inc | 0.0016 | 0.0003 | -0.0014 | 0.6124 | 0.8268 | 0.2144 | 0.0007 | 0.0002 | -0.0005 |
| Comerica Incorporated | 0.0027 | 0.0020 | -0.0007 | 0.5866 | 0.6768 | 0.0902 | 0.0005 | 0.0001 | -0.0005 |
| Fifth Third Bancorp | 0.0027 | 0.0031 | 0.0004 | 0.5620 | 0.6085 | 0.0465 | 0.0015 | 0.0001 | -0.0014 |
| Jpmorgan Chase & Co. | 0.0016 | 0.0023 | 0.0007 | 0.6890 | 0.6478 | -0.0412 | 0.0014 | 0.0002 | -0.0012 |
| Keycorp | 0.0010 | 0.0024 | 0.0014 | 0.6773 | 0.6841 | 0.0068 | 0.0012 | -0.0001 | -0.0013 |
| M&T Bank Corporation | 0.0030 | 0.0030 | 0.0000 | 0.5517 | 0.5859 | 0.0341 | 0.0004 | 0.0002 | -0.0002 |
| Northern Trust Corporation | 0.0030 | 0.0020 | -0.0010 | 0.6649 | 0.7184 | 0.0535 | 0.0002 | -0.0001 | -0.0002 |
| Pnc Financial Services Group, Inc., The | 0.0034 | 0.0030 | -0.0004 | 0.6318 | 0.6291 | -0.0027 | 0.0011 | 0.0001 | -0.0010 |
| State Street Corporation | 0.0019 | 0.0024 | 0.0005 | 0.7395 | 0.7507 | 0.0112 | 0.0001 | 0.0000 | -0.0001 |
| Suntrust Banks, Inc. | 0.0020 | 0.0021 | 0.0001 | 0.6571 | 0.6953 | 0.0382 | 0.0008 | 0.0002 | -0.0006 |
| U.S. Bancorp | 0.0043 | 0.0037 | -0.0005 | 0.4811 | 0.5378 | 0.0567 | 0.0008 | 0.0003 | -0.0005 |
| Wells Fargo & Company | 0.0035 | 0.0033 | -0.0002 | 0.5871 | 0.5985 | 0.0115 | 0.0011 | 0.0002 | -0.0009 |
| Zions Bancorporation | 0.0014 | 0.0014 | 0.0000 | 0.6416 | 0.7249 | 0.0833 | 0.0008 | -0.0001 | -0.0009 |

Table 2.8: Weights. Big representative bank

| Controls | Trading Assets | IB advisory, brokerage, and underwriting | Derivatives held for trading | Derivatives held for non-trading | Z-score | AFS securities |
|---|---|---|---|---|---|---|
| 1St Source | | 0.027 | | 0.000 | | 0.000 |
| American National Bankshares | 0.008 | 0.048 | 0.000 | 0.100 | 0.028 | 0.032 |
| Associated Banc-Corp | 0.000 | 0.000 | -0.109 | -0.103 | -0.036 | 0.018 |
| Bancfirst | 0.051 | 0.056 | | 0.044 | 0.000 | -0.009 |
| Bank Of Hawaii | 0.005 | 0.032 | 0.132 | 0.068 | 0.000 | 0.088 |
| Bar Harbor Bankshares | | 0.020 | | 0.143 | 0.007 | 0.078 |
| Bb&T | 0.078 | 0.036 | 0.057 | 0.048 | -0.026 | 0.031 |
| C&F Financial | | 0.137 | | 0.107 | 0.078 | 0.000 |
| Capital City Bank Group, | | 0.000 | | | 0.000 | -0.048 |
| Center Bancorp, | | -0.156 | | | 0.000 | -0.079 |
| Central Pacific Financial Corp. | | -0.089 | | 0.000 | -0.006 | 0.000 |
| Century Bancorp, | | -0.029 | | | -0.067 | -0.061 |
| City Holding Company | | 0.024 | | 0.018 | 0.000 | 0.000 |
| Columbia Banking System, | 0.076 | 0.000 | | 0.000 | 0.073 | 0.000 |
| Commerce Bancshares, | 0.000 | 0.015 | 0.142 | 0.021 | 0.008 | 0.255 |
| Cullen/Frost Bankers, | 0.032 | 0.000 | 0.000 | 0.087 | 0.081 | 0.100 |
| Cvb Financial Corp. | -0.029 | -0.079 | | -0.161 | -0.044 | -0.057 |
| Enterprise Bancorp, | 0.025 | -0.036 | | | -0.015 | -0.078 |
| Enterprise Financial Services Corp | | 0.055 | | 0.000 | 0.000 | 0.026 |
| Farmers Capital Bank | | -0.017 | | | 0.028 | -0.015 |
| Farmers National Banc Corp. | | -0.043 | | | -0.087 | -0.045 |
| Financial Institutions, | | -0.094 | | -0.070 | -0.099 | -0.052 |
| First Bancorp | | -0.019 | | -0.029 | -0.043 | 0.020 |
| First Busey | | 0.000 | | | 0.018 | -0.027 |
| First Citizens Banc Corp | | -0.018 | | | -0.080 | 0.046 |
| First Citizens Bancshares, | | 0.000 | | 0.000 | 0.082 | 0.000 |
| First Commonwealth Financial | | 0.013 | | 0.038 | 0.014 | -0.018 |
| First Community Bancshares, | | 0.053 | | 0.031 | 0.015 | 0.000 |
| First Financial | | 0.077 | | 0.044 | 0.090 | -0.047 |
| First Merchants | | 0.000 | | 0.058 | 0.000 | 0.024 |
| First Mid Ill Bancshares Inc | | -0.116 | | | -0.106 | -0.106 |
| First Midwest Bancorp, | 0.000 | 0.003 | 0.000 | | -0.073 | 0.000 |
| Fulton Financial | | 0.215 | | 0.182 | 0.178 | 0.000 |
| German American Bancorp, | | -0.071 | | -0.084 | -0.092 | -0.009 |
| Hanmi Financial | | 0.230 | | 0.152 | 0.176 | 0.108 |
| Heartland Financial Usa, | | 0.000 | 0.080 | 0.000 | 0.000 | 0.154 |
| Heritage Commerce Corp | | | | | | 0.064 |
| Home Bancshares, | 0.017 | 0.016 | | | 0.000 | -0.043 |

Table 2.9: Weights. Big representative bank. Continued

| Controls | Trading Assets | IB advisory, brokerage, and underwriting | Derivatives held for trading | Derivatives held for non-trading | Z-score | AFS securities |
|---|---|---|---|---|---|---|
| Huntington Bancshares Incorporated | 0.000 | -0.130 | 0.000 | -0.095 | -0.060 | 0.003 |
| Iberiabank | | -0.030 | | 0.000 | -0.007 | -0.003 |
| Independent Bank | 0.028 | 0.211 | -0.059 | 0.341 | 0.299 | 0.197 |
| Independent Bank Corp. | | 0.000 | 0.313 | -0.015 | 0.000 | 0.000 |
| International Bancshares | | 0.022 | | -0.033 | 0.031 | -0.021 |
| Lakeland Bancorp, | | 0.083 | | | 0.074 | 0.063 |
| Lakeland Financial | | 0.046 | | | | 0.000 |
| Midsouth Bancorp, | | 0.040 | | | 0.061 | 0.147 |
| Midwestone Financial Group, | | 0.014 | | -0.077 | 0.030 | -0.020 |
| Nara Bancorp, | | 0.054 | | 0.010 | 0.032 | 0.093 |
| Nbt Bancorp | | 0.025 | -0.031 | | 0.063 | -0.087 |
| Park National | 0.084 | 0.146 | | 0.183 | 0.117 | 0.006 |
| Peoples Bancorp | | -0.011 | | -0.087 | -0.042 | -0.038 |
| Peoples Utah Bancorp | | 0.003 | | | 0.034 | 0.000 |
| Popular, | 0.200 | 0.176 | | 0.241 | 0.171 | 0.277 |
| Prosperity Bancshares, | | 0.000 | | | 0.066 | 0.000 |
| Qcr Holdings, | | 0.000 | | 0.000 | | 0.035 |
| Regions Financial | 0.379 | 0.162 | 0.197 | 0.180 | 0.191 | 0.182 |
| Renasant | | -0.136 | | -0.016 | -0.044 | -0.165 |
| Republic Bancorp, | 0.090 | 0.000 | 0.060 | 0.000 | 0.054 | 0.071 |
| Royal Bancshares Of Pennsylvania, | | -0.012 | | 0.000 | -0.021 | -0.025 |
| S & T Bancorp, | 0.001 | -0.099 | -0.015 | -0.055 | -0.051 | -0.203 |
| S. Y. Bancorp, | 0.049 | 0.106 | | | 0.041 | 0.000 |
| Sandy Spring Bancorp, | 0.042 | 0.077 | | | 0.003 | 0.084 |
| Scbt Financial | | 0.000 | | 0.000 | -0.031 | -0.047 |
| Shore Bancshares, | | -0.183 | | -0.100 | -0.021 | -0.216 |
| Southwest Georgia Financial | | 0.000 | | | -0.053 | 0.000 |
| Summit Financial Group, | | 0.033 | | 0.001 | 0.034 | 0.000 |
| Texas Capital Bancshares, | | 0.004 | | | 0.048 | -0.019 |
| Tompkins Financial | 0.032 | 0.076 | 0.019 | 0.040 | 0.020 | 0.129 |
| Trustmark | 0.040 | 0.045 | 0.000 | 0.000 | 0.049 | 0.011 |
| Union First Market Bankshares | | 0.055 | | | -0.002 | 0.085 |
| United Community Banks, | 0.000 | -0.063 | | -0.033 | -0.079 | -0.037 |
| Valley National Bancorp | | 0.026 | 0.259 | 0.010 | 0.091 | 0.193 |
| West Ban, | | 0.000 | | 0.022 | -0.076 | 0.000 |
| **Sum of Weights** | 1.210 | 1.028 | 1.046 | 1.214 | 1.124 | 1.046 |

Table 2.10: Weights. Small representative bank

| Controls | Trading Assets | IB advisory, brokerage, and underwriting | Derivatives held for trading | Derivatives held for non-trading | Z-score | AFS securities |
|---|---|---|---|---|---|---|
| 1St Source | 0 | 0.000 | 0.000 | -0.031 | | 0.040 |
| American National Bankshares | 0.000 | 0.000 | 0.000 | 0.000 | 0.003 | 0.053 |
| Associated Banc-Corp | 0.000 | 0.000 | | 0.000 | 0.002 | 0.026 |
| Bancfirst | | 0.009 | | 0.002 | 0.009 | 0.000 |
| Bank Of Hawaii | 0.000 | 0.000 | | 0.000 | 0.000 | 0.000 |
| Bar Harbor Bankshares | | 0.000 | 0.000 | -0.074 | 0.000 | 0.006 |
| Bb&T | 0.071 | 0.017 | 0.183 | 0.000 | 0.025 | 0.001 |
| C&F Financial | | 0.000 | | 0.014 | 0.000 | 0.079 |
| Capital City Bank Group, | | 0.070 | | | 0.000 | 0.000 |
| Center Bancorp, | | 0.000 | | | -0.030 | 0.000 |
| Central Pacific Financial Corp. | | 0.014 | | -0.054 | 0.014 | 0.082 |
| Century Bancorp, | | 0.034 | | | -0.006 | 0.000 |
| City Holding Company | | 0.000 | | 0.015 | 0.000 | 0.000 |
| Columbia Banking System, | 0.013 | 0.000 | | 0.077 | 0.038 | 0.000 |
| Commerce Bancshares, | 0.045 | 0.000 | 0.075 | 0.066 | 0.067 | 0.079 |
| Cullen/Frost Bankers, | 0.000 | 0.000 | 0.043 | -0.074 | 0.000 | 0.011 |
| Cvb Financial Corp. | 0.000 | 0.000 | | -0.066 | 0.000 | 0.014 |
| Enterprise Bancorp, | 0.000 | 0.018 | | | 0.000 | 0.021 |
| Enterprise Financial Services Corp | | 0.000 | | -0.022 | 0.000 | 0.000 |
| Farmers Capital Bank | | 0.000 | | | 0.028 | 0.065 |
| Farmers National Banc Corp. | | 0.000 | | | -0.031 | -0.031 |
| Financial Institutions, | | 0.049 | | -0.019 | 0.000 | 0.000 |
| First Bancorp | | 0.200 | | 0.086 | 0.016 | 0.001 |
| First Busey | | 0.000 | | | 0.048 | 0.021 |
| First Citizens Banc Corp | | 0.000 | | | 0.022 | 0.007 |
| First Citizens Bancshares, | | 0.000 | | 0.061 | 0.047 | 0.039 |
| First Commonwealth Financial | | 0.000 | | 0.000 | 0.000 | 0.000 |
| First Community Bancshares, | | 0.000 | | 0.023 | 0.000 | -0.022 |
| First Financial | | 0.030 | | 0.021 | 0.019 | -0.007 |
| First Merchants | | 0.070 | | 0.000 | 0.091 | 0.000 |
| First Mid Ill Bancshares Inc | | 0.020 | | | 0.000 | -0.056 |
| First Midwest Bancorp, | -0.061 | 0.043 | 0.000 | | 0.000 | -0.115 |
| Fulton Financial | | 0.030 | | 0.000 | 0.142 | 0.121 |
| German American Bancorp, | | 0.075 | | 0.000 | -0.002 | -0.003 |
| Hanmi Financial | | 0.002 | | 0.154 | 0.042 | 0.201 |
| Heartland Financial Usa, | | 0.122 | -0.112 | -0.191 | 0.000 | 0.010 |
| Heritage Commerce Corp | | | | | | 0.025 |
| Home Bancshares, | -0.024 | 0.000 | | | -0.031 | 0.000 |

Table 2.11: Weights. Small representative bank. Continued

| Controls | Trading Assets | IB advisory, brokerage, and underwriting | Derivatives held for trading | Derivatives held for non-trading | Z-score | AFS securities |
|---|---|---|---|---|---|---|
| Huntington Bancshares Incorporated | 0.000 | 0.072 | 0.031 | -0.049 | 0.000 | -0.027 |
| Iberiabank | | 0.000 | | 0.000 | 0.047 | 0.019 |
| Independent Bank | | 0.000 | 0.136 | 0.406 | 0.064 | 0.040 |
| Independent Bank Corp. | 0.003 | 0.000 | 0.349 | 0.031 | 0.015 | 0.049 |
| International Bancshares | | 0.058 | | 0.112 | 0.027 | 0.008 |
| Lakeland Bancorp, | | -0.001 | | | 0.000 | 0.000 |
| Lakeland Financial | | 0.058 | | | | 0.018 |
| Midsouth Bancorp, | | 0.000 | | | 0.030 | 0.077 |
| Midwestone Financial Group, | | 0.000 | | 0.000 | 0.069 | 0.054 |
| Nara Bancorp, | | 0.000 | | 0.062 | 0.034 | 0.000 |
| Nbt Bancorp | | 0.021 | 0.000 | | 0.000 | 0.000 |
| Park National | 0.125 | 0.051 | | 0.048 | 0.066 | 0.004 |
| Peoples Bancorp | | 0.197 | | 0.204 | 0.020 | 0.017 |
| Peoples Utah Bancorp | | 0.000 | | | 0.019 | 0.048 |
| Popular, | 0.151 | 0.000 | | 0.065 | 0.101 | 0.020 |
| Prosperity Bancshares, | | -0.023 | | | 0.000 | 0.000 |
| Qcr Holdings, | | 0.000 | | -0.018 | | 0.000 |
| Regions Financial | 0.388 | 0.001 | 0.000 | 0.081 | 0.046 | 0.128 |
| Renasant | | 0.000 | | -0.043 | -0.024 | -0.020 |
| Republic Bancorp, | 0.005 | 0.030 | -0.022 | -0.105 | 0.000 | -0.009 |
| Royal Bancshares Of Pennsylvania, | | 0.000 | | 0.056 | 0.000 | -0.006 |
| S & T Bancorp, | 0.277 | 0.000 | 0.027 | 0.000 | 0.040 | 0.050 |
| S. Y. Bancorp, | 0.051 | 0.000 | | | 0.034 | 0.000 |
| Sandy Spring Bancorp, | 0.052 | -0.020 | | | 0.000 | 0.000 |
| Scbt Financial | | -0.018 | | -0.026 | -0.025 | 0.000 |
| Shore Bancshares, | | 0.000 | | 0.000 | 0.064 | 0.038 |
| Southwest Georgia Financial | | 0.010 | | | 0.035 | 0.056 |
| Summit Financial Group, | | 0.000 | | 0.000 | 0.000 | -0.015 |
| Texas Capital Bancshares, | | 0.000 | | | 0.000 | 0.000 |
| Tompkins Financial | 0.070 | 0.000 | 0.062 | -0.005 | 0.000 | 0.020 |
| Trustmark | 0.033 | 0.016 | 0.012 | 0.031 | 0.000 | 0.000 |
| Union First Market Bankshares | | 0.000 | | | 0.039 | 0.000 |
| United Community Banks, | 0.000 | 0.028 | | 0.082 | 0.006 | 0.000 |
| Valley National Bancorp | | 0.000 | 0.237 | 0.004 | 0.066 | 0.073 |
| West Ban, | | 0.020 | | 0.170 | 0.116 | 0.184 |
| **Sum of Weights** | 1.199 | 1.301 | 1.019 | 1.094 | 1.397 | 1.494 |

Table 2.12: stimates of gaps for a representative treated unit. Big Banks. Calculated as observed minus predicted

| Variable of interest | | | Controlling variables | | |
|---|---|---|---|---|---|
| | Pre-Final Rule | Post-Final Rule | Profitability | Efficiency | Asset Quality |
| Trading Assets | -0.0075 | -0.0175 | -0.0002 | -0.0020 | 0.00021 |
| | (0.28) | (0.16) | (0.44) | (0.55) | (0.16) |
| IB advisory, brokerage, and underwriting | 0.024 | 0.071 | 0.0005 | -0.031 | 0.0001 |
| | (0.53) | (0.29) | (0.458) | (0.25) | (0.10) |
| Total GNA of derivative contracts held for trading | -4.9E+09 | -8.5E+09 | 0.00024 | 0.044 | 0.00022 |
| | (0.542) | (0.417) | (0.79) | (0.25) | (0.75) |
| Total GNA of derivative contracts held for purpose other than trading | 4.8E+08** | 2.1E+08 | 0.00033 | -0.0661 | 0.00024 |
| | (0.019) | (0.26) | (0.38) | (0.12) | (0.38) |
| Z-score | 80.5 | 124.8 | 0.0003 | -0.0228 | 0.0004 |
| | (0.47) | (0.45) | (0.565) | (0.377) | (0.71) |
| Available for sale securities | 0.0096 | 0.0054 | 0.00044 | -0.041 | -5.0E-05 |
| | (0.84) | (0.93) | (0.42) | (0.21) | (0.28) |

P-values are in parenthesis; *** p<0.01, ** p<0.05

Table 2.13: Estimates of gaps for a representative treated unit. Small Banks. Calculated as observed minus predicted

| Variable of interest | | | Controlling variables | | |
|---|---|---|---|---|---|
| | Pre-Final Rule | Post-Final Rule | Profitability | Efficiency | Asset Quality |
| Trading Assets | 0.0003 | -0.0003 | -0.001 | 0.0019 | 0.0019 |
| | (0.78) | (0.89) | (0.06) | (0.33) | (0.33) |
| IB advisory, brokerage, and underwriting | -5.85E-05 | -1.21E-03 | -0.0005 | 0.050 | 0.00016 |
| | (0.764) | (0.736) | (0.417) | (0.736) | (0.764) |
| Total GNA of derivative contracts held for trading | -621580.7 | 2499686.2 | -0.00082 | 0.074 | 0.000 |
| | (0.625) | (0.375) | (0.167) | (0.21) | (0.66) |
| Total GNA of derivative contracts held for purpose other than trading | -1125711 | -1928116 | -0.00013 | 0.06 | 7.7E-05 |
| | (0.7) | (0.7) | (0.32) | (0.64) | (0.84) |
| Z-score | 9.3 | 20.4 | -0.0007 | 0.0573 | 0.00001 |
| | (0.99) | (0.86) | (0.275) | (0.71) | (0.579) |
| Available for sale securities | -0.004 | -0.021 | -0.0005 | 0.0357 | 6.8E-0.5 |
| | (0.99) | (0.726) | (0.53) | (0.93) | (0.59) |

P-values are in parenthesis; *** p<0.01, ** p<0.05

Table 2.14: Estimates of gaps in Trading Assets as the percentage of Total Assets. Big Banks. Calculated as observed minus predicted. Bank by bank calculation

| Banks | Pre-Final Rule Avg Gap | Pre-Final Rule P-value | Post-Final Rule Avg Gap | Post-Final Rule P-value |
|---|---|---|---|---|
| Bank Of America Corporation | -0.0004 | 0.39 | 0.0012 | 0.56 |
| Citigroup Inc. | -0.0126 | 0.50 | -0.0256 | 0.33 |
| Citizens Financial Group Inc | 0.0034 | 0.22 | 0.0022 | 0.44 |
| Comerica Incorporated | 0.0006 | 0.50 | -0.0016 | 0.56 |
| Fifth Third Bancorp | 0.0049 | 0.44 | 0.0088 | 0.33 |
| Jpmorgan Chase & Co. | -0.0399 | 0.22 | -0.0745 | 0.11 |
| Keycorp | -0.0075 | 0.22 | -0.0074 | 0.33 |
| M&T Bank Corporation | -0.0002 | 0.83 | -0.0027 | 0.22 |
| Northern Trust Corporation | -0.0028 | 0.28 | -0.0085 | 0.33 |
| Pnc Financial Services Group, Inc., The | -0.0129 | 0.22 | -0.0181 | 0.22 |
| State Street Corporation | -0.0055 | 0.28 | -0.0132 | 0.33 |
| Suntrust Banks, Inc. | 0.0062 | 0.56 | 0.0139 | 0.33 |
| U.S. Bancorp | 0.0013 | 0.39 | 0.0036 * | 0.06 |
| Wells Fargo & Company | 0.0186 * | 0.06 | 0.0266 * | 0.06 |
| Zions Bancorporation | 0.0001 | 0.50 | 0.0063 | 0.33 |

Table 2.15: Estimates of gaps in Trading Assets as the percentage of Total Assets. Small Banks. Calculated as observed minus predicted. Bank by bank calculation

| Banks | Pre-Final Rule Avg Gap | Pre-Final Rule P-value | Post-Final Rule Avg Gap | Post-Final Rule P-value |
|---|---|---|---|---|
| Banner Corporation | 0.0077 | 0.50 | 0.0062 | 0.78 |
| Bok Financial Corporation | -0.0005 | 0.89 | -0.0013 | 0.83 |
| Boston Private Financial Holdings, Inc. | 0.0001 | 0.39 | 0.0013 * | 0.06 |
| Camden National Corporation | 0.0002 | 0.17 | 0.0005 *** | 0.00 |
| Cathay General Bancorp | -0.0014 | 0.28 | -0.0024 | 0.72 |
| Community Bank System, Inc. | 0 | 0.94 | 0 | 0.78 |
| First Financial Bancorp | -0.0001 | 0.94 | -0.0001 | 0.94 |
| First Financial Bankshares, Inc. | -0.0045 | 0.50 | -0.0077 | 0.33 |
| First Horizon National Corporation | 0.0083 | 0.39 | 0.0048 | 0.56 |
| First United Corporation | -0.0089 | 0.50 | -0.0059 | 0.83 |
| Seacoast Banking Corporation Of Florida | -0.0033 | 0.72 | -0.0026 | 0.89 |
| Svb Financial Group | 0.0007 | 0.28 | 0.0018 | 0.17 |
| Synovus Financial Corp. | 0.001 | 0.39 | 0.0007 | 0.72 |
| Umb Financial Corporation | -0.0021 | 0.28 | -0.0039 | 0.22 |
| Wesbanco, Inc. | 0.0004 *** | 0.00 | 0.0006 *** | 0.00 |
| Wintrust Financial Corporation | 0.0006 | 0.11 | 0.0019 * | 0.06 |

Table 2.16: Estimates of gaps in IB advisory, brockerage, and underwriting fees and commission as a percentage of interest and non-interest income. Big Banks. Calculated as observed minus predicted. Bank by bank calculation

| Banks | Pre-Final Rule Avg Gap | Pre-Final Rule P-value | Post-Final Rule Avg Gap | Post-Final Rule P-value |
|---|---|---|---|---|
| Bank Of America Corporation | 0.055 | 0.17 | 0.1974 ** | 0.04 |
| Citigroup Inc. | -0.0469 | 0.90 | 0.0062 | 1.00 |
| Citizens Financial Group Inc | 0.0078 | 0.14 | 0.0114 | 0.18 |
| Comerica Incorporated | -0.0037 | 0.68 | -0.0022 | 0.85 |
| Fifth Third Bancorp | 0.0024 | 0.17 | 0.0066 | 0.72 |
| Jpmorgan Chase & Co. | -0.0044 | 0.35 | 0.0026 | 0.76 |
| Keycorp | -0.0424 | 0.20 | 0.0117 | 0.63 |
| M&T Bank Corporation | -0.0184 * | 0.08 | -0.0091 | 0.48 |
| Northern Trust Corporation | -0.0001 | 0.34 | -0.0003 | 0.25 |
| Pnc Financial Services Group, Inc., The | 0.0076 | 0.96 | -0.011 | 0.97 |
| State Street Corporation | -0.014 | 0.23 | -0.001 | 0.92 |
| Suntrust Banks, Inc. | 0.0064 | 0.63 | 0.048 * | 0.06 |
| U.S. Bancorp | -0.0027 | 0.70 | 0.0093 | 0.39 |
| Wells Fargo & Company | 0.0606 * | 0.07 | 0.1052 ** | 0.04 |
| Zions Bancorporation | -0.0029 | 0.15 | -0.0014 | 0.73 |

Table 2.17: Estimates of gaps in IB advisory, brockerage, and underwriting fees and commission as a percentage of interest and non-interest income. Small Banks. Calculated as observed minus predicted. Bank by bank calculation

| Banks | Pre-Final Rule Avg Gap | Pre-Final Rule P-value | Post-Final Rule Avg Gap | Post-Final Rule P-value |
|---|---|---|---|---|
| Auburn National Bancorporation, Inc. | -0.0004 | 0.24 | -0.0006 | 0.25 |
| Banner Corporation | -0.0002 | 0.14 | 0.0011 ** | 0.03 |
| Bok Financial Corporation | 0.0008 | 0.75 | 0.0152 | 0.14 |
| Boston Private Financial Holdings, Inc. | 0.0622 | 0.79 | 0.0287 | 1.00 |
| Camden National Corporation | -0.0101 | 0.46 | -0.0222 | 0.18 |
| Cathay General Bancorp | 0.009 ** | 0.04 | 0.0045 * | 0.06 |
| Chemung Financial Corp | 0.0035 | 0.21 | 0.0099 ** | 0.03 |
| Colony Bankcorp, Inc. | -0.0004 | 0.27 | -0.0006 | 0.45 |
| Community Bank System, Inc. | -0.0023 | 0.49 | 0.0418 *** | 0.00 |
| F.N.B. Corporation | 0.0073 | 0.44 | -0.0002 | 0.94 |
| First Financial Bancorp | -0.0128 | 0.99 | -0.0146 | 1.00 |
| First Financial Bankshares, Inc. | 0.0002 | 0.62 | 0.0018 | 0.23 |
| First Horizon National Corporation | -0.0477 | 0.17 | -0.0167 | 0.93 |
| First United Corporation | -0.013 * | 0.10 | -0.0547 ** | 0.03 |
| Heritage Financial Corporation | 6e-04 * | 0.07 | 0.0031 *** | 0.00 |
| Mainsource Financial Group, Inc. | -0.0041 | 0.68 | -0.0064 | 0.87 |
| Old National Bancorp | -0.0037 | 0.68 | -0.0053 | 0.94 |
| Old Second Bancorp, Inc. | 0.0033 | 0.13 | 0.0013 | 0.72 |
| Rockville Financial, Inc. | -0.0035 | 0.35 | -0.0075 | 0.32 |
| Seacoast Banking Corporation Of Florida | -0.0009 | 0.42 | -0.0068 | 0.45 |
| Southside Bancshares, Incorporated | -0.0024 | 0.25 | 0.0008 | 0.35 |
| Sun Bancorp, Inc | -0.0006 | 0.65 | -0.0059 | 0.72 |
| Svb Financial Group | -0.0364 | 0.55 | -0.0583 | 0.41 |
| Synovus Financial Corp. | 0.0209 | 0.69 | 0.0195 | 0.94 |
| Trico Bancshares | 0.0053 | 0.44 | 0.0064 | 0.56 |
| Umb Financial Corporation | 0.0019 | 0.55 | -0.0019 | 0.55 |
| United Bankshares, Inc. | 0.0028 | 0.55 | 0.0123 | 0.15 |
| Washington Trust Bancorp, Inc. | -0.0193 | 0.70 | 0.0137 | 0.65 |
| Wesbanco, Inc. | 0.0073 | 0.55 | 0.0061 | 0.55 |
| Wintrust Financial Corporation | 0.0039 | 0.20 | -0.0437 | 0.18 |

Table 2.18: Estimates of gaps in derivatives held for trading. Big Banks. Calculated as observed minus predicted. Bank by bank calculation

| Banks | Pre-Final Rule Avg Gap | Pre-Final Rule P-value | Post-Final Rule Avg Gap | Post-Final Rule P-value |
|---|---|---|---|---|
| Bank Of America Corporation | -3426424401 | 0.67 | -28201411417 | 0.33 |
| Citigroup Inc. | 14271902067 | 0.33 | 20237822883 | 0.33 |
| Citizens Financial Group Inc | -8188296 | 0.42 | 16988034 | 0.33 |
| Comerica Incorporated | 2226795 | 0.54 | 2535154 | 0.63 |
| Fifth Third Bancorp | -8438739 | 0.42 | 6863972 | 0.63 |
| Jpmorgan Chase & Co. | -19801069324 | 0.54 | -45256572509 | 0.33 |
| Keycorp | -31029728 * | 0.08 | -30491602 | 0.21 |
| M&T Bank Corporation | -16845 | 0.54 | -2357333 | 0.63 |
| Northern Trust Corporation | 73398256 | 0.50 | 60217546 | 0.63 |
| Pnc Financial Services Group, Inc., The | -24079144 | 0.79 | 18678976 | 0.92 |
| State Street Corporation | 481741486 | 0.25 | 714682435 | 0.25 |
| Suntrust Banks, Inc. | -10204762 | 0.67 | -22424581 | 0.67 |
| U.S. Bancorp | -11039624 | 0.38 | 83224620 ** | 0.04 |
| Wells Fargo & Company | -464798393 | 0.54 | 2159546004 | 0.38 |
| Zions Bancorporation | -927494 | 0.38 | -743712 | 0.67 |

Table 2.19: Estimates of gaps in derivatives held for trading. Small Banks. Calculated as observed minus predicted. Bank by bank calculation

| Banks | Pre-Final Rule Avg Gap | Pre-Final Rule P-value | Post-Final Rule Avg Gap | Post-Final Rule P-value |
|---|---|---|---|---|
| Bok Financial Corporation | 6373137 | 0.58 | 12491934 | 0.58 |
| Boston Private Financial Holdings, Inc. | -79580 | 0.42 | 343498* | 0.08 |
| Cathay General Bancorp | 123771** | 0.04 | 24477 | 0.13 |
| First Horizon National Corporation | -2726855 | 0.63 | -3199752 | 0.71 |
| Investors Bancorp, Mhc | -59782 * | 0.08 | -92623 | 0.21 |
| Sun Bancorp, Inc | -971172 | 0.33 | -1485928 | 0.21 |
| Svb Financial Group | 325065 | 0.42 | 2162472** | 0.04 |
| Synovus Financial Corp. | -2641009 * | 0.08 | -1646348 | 0.42 |
| Umb Financial Corporation | 18772 | 0.46 | 18126 | 0.58 |
| Wintrust Financial Corporation | 1945631*** | 0.00 | 5041557*** | 0.00 |

Table 2.20: Estimates of gaps in derivatives held for the purpose other than trading. Big Banks. Calculated as observed minus predicted. Bank by bank calculation

| Banks | Pre-Final Rule Avg Gap | Pre-Final Rule P-value | Post-Final Rule Avg Gap | Post-Final Rule P-value |
|---|---|---|---|---|
| Bank Of America Corporation | 1561979334*** | 0.00 | 655931303 | 0.37 |
| Citigroup Inc. | -547824150 | 0.31 | -1056591437 | 0.27 |
| Citizens Financial Group Inc | -3997983 | 0.73 | -21910976 | 0.51 |
| Comerica Incorporated | -4505760 | 0.65 | -1464127 | 0.98 |
| Fifth Third Bancorp | 2066138 | 0.71 | -3731452 | 0.86 |
| Jpmorgan Chase & Co. | 623370352** | 0.02 | 485263223 | 0.29 |
| Keycorp | 2301602 | 0.84 | -2111189 | 0.92 |
| M&T Bank Corporation | 266176 | 0.78 | 781776 | 0.71 |
| Northern Trust Corporation | 4609215 | 0.12 | 1904707 | 0.71 |
| Pnc Financial Services Group, Inc., The | 153453048 | 0.12 | 86652714 | 0.63 |
| State Street Corporation | 3394965 | 0.20 | 16271787*** | 0.00 |
| Suntrust Banks, Inc. | 168563 | 0.98 | -8160482 | 1.00 |
| U.S. Bancorp | 1658146 | 0.86 | 5575592 | 0.84 |
| Wells Fargo & Company | -156174479 | 0.86 | -174962154 | 0.86 |
| Zions Bancorporation | -2172486 | 0.27 | -9641837 | 0.29 |

Table 2.21: Estimates of gaps in derivatives held for the purpose other than trading. Small Banks. Calculated as observed minus predicted. Bank by bank calculation

| Banks | Pre-Final Rule Avg Gap | Pre-Final Rule P-value | Post-Final Rule Avg Gap | Post-Final Rule P-value |
|---|---|---|---|---|
| Auburn National Bancorporation, Inc. | -8165 | 0.57 | -15810 | 0.55 |
| Banner Corporation | 9701 | 0.98 | 70234 | 0.63 |
| Bok Financial Corporation | -1328 | 1.00 | 818393 | 0.31 |
| Boston Private Financial Holdings, Inc. | 42684 | 0.55 | 111379 | 0.51 |
| Camden National Corporation | -103712 | 0.37 | 98329 | 0.39 |
| Cathay General Bancorp | -155986 * | 0.08 | 200243 | 0.29 |
| Community Bank System, Inc. | -64034 | 0.35 | -173764 | 0.22 |
| F.N.B. Corporation | -283188 | 0.18 | -78698 | 0.82 |
| First Financial Bancorp | 631182 | 0.12 | 859981 | 0.29 |
| First Financial Bankshares, Inc. | 18300 | 0.51 | 36316 | 0.39 |
| First Horizon National Corporation | -11100739 | 0.57 | -27883061 | 0.51 |
| Old National Bancorp | 422165 | 0.78 | 73036 | 0.92 |
| Old Second Bancorp, Inc. | -145970 | 0.41 | -208172 | 0.53 |
| Rockville Financial, Inc. | 74228*** | 0.00 | 820985*** | 0.00 |
| Seacoast Banking Corporation Of Florida | 24579 | 0.84 | 86302 | 0.57 |
| Southside Bancshares, Incorporated | 89076*** | 0.00 | 128151*** | 0.00 |
| Sun Bancorp, Inc | 455088 | 0.55 | -683275 | 0.69 |
| Svb Financial Group | -361419 | 0.71 | -1365489 | 0.41 |
| Synovus Financial Corp. | -706213 | 0.84 | -1344045 | 0.78 |
| Umb Financial Corporation | 106687** | 0.02 | 564896*** | 0.00 |
| United Bankshares, Inc. | -460591 | 0.55 | -740773 | 0.51 |
| Washington Trust Bancorp, Inc. | 6243 | 0.78 | 100030 | 0.41 |
| Wesbanco, Inc. | -25839 | 0.61 | 242892** | 0.04 |
| Wintrust Financial Corporation | -149326 | 0.57 | -805052 | 0.57 |

Table 2.22: Estimates of gaps in Z-score. Big Banks. Calculated as observed minus predicted. Bank by bank calculation

| Banks | Pre-Final Rule Avg Gap | Pre-Final Rule P-value | Post-Final Rule Avg Gap | Post-Final Rule P-value |
|---|---|---|---|---|
| Bank Of America Corporation | -52 | 0.94 | -104 | 0.86 |
| Citigroup Inc. | 69 | 0.12 | 160*** | 0.00 |
| Citizens Financial Group Inc | 62 | 0.57 | -45 | 0.42 |
| Comerica Incorporated | 97 | 0.13 | 216** | 0.03 |
| Fifth Third Bancorp | 67 | 0.58 | -16 | 0.96 |
| Jpmorgan Chase & Co. | 54 | 0.71 | 130* | 0.10 |
| Keycorp | 177 | 0.30 | 105 | 0.39 |
| M&T Bank Corporation | -658 | 0.32 | -690 | 0.38 |
| Northern Trust Corporation | 241 | 0.42 | 224 | 0.52 |
| Pnc Financial Services Group, Inc., The | -21 | 1.00 | 482** | 0.01 |
| State Street Corporation | -28 | 0.74 | -90 | 0.62 |
| Suntrust Banks, Inc. | 179 | 0.71 | -213 | 0.81 |
| U.S. Bancorp | -714 | 0.39 | -466 | 0.41 |
| Wells Fargo & Company | 77 | 0.78 | 120 | 0.70 |
| Zions Bancorporation | 82 | 0.93 | -9 | 1.00 |

Table 2.23: Estimates of gaps in Zscore. Small Banks. Calculated as observed minus predicted. Bank by bank calculation

| Banks | Pre-Final Rule Avg Gap | Pre-Final Rule P-value | Post-Final Rule Avg Gap | Post-Final Rule P-value |
|---|---|---|---|---|
| Auburn National Bancorporation, Inc. | 400 | 0.68 | 451 | 0.70 |
| Banner Corporation | -145 | 0.86 | -194 | 0.88 |
| Bok Financial Corporation | -115 | 0.81 | -555 | 0.16 |
| Boston Private Financial Holdings, Inc. | 40 | 0.74 | 147 | 0.86 |
| Camden National Corporation | -123 | 0.55 | -392 | 0.33 |
| Cathay General Bancorp | 348** | 0.03 | 262 | 0.13 |
| Chemung Financial Corp | -21 | 0.99 | -117 | 0.84 |
| Colony Bankcorp, Inc. | -208 | 0.61 | -660 | 0.77 |
| Community Bank System, Inc. | 43 | 0.80 | 145 | 0.88 |
| F.N.B. Corporation | 456 | 0.48 | 703 | 0.26 |
| First Financial Bancorp | 237 | 0.55 | 299 | 0.77 |
| First Financial Bankshares, Inc. | -106 | 0.84 | 217 | 0.72 |
| First Horizon National Corporation | 25 | 0.91 | -182 | 0.46 |
| First United Corporation | -100 | 0.49 | -73 | 0.43 |
| Heritage Financial Corporation | -382 | 0.35 | -325 | 0.39 |
| Investors Bancorp, Mhc | 254 | 0.28 | 340 | 0.13 |
| Mainsource Financial Group, Inc. | -146 | 0.74 | -651 | 0.22 |
| Old National Bancorp | 124** | 0.04 | 177 | 0.10 |
| Old Second Bancorp, Inc. | -195 | 0.57 | -356 | 0.42 |
| Pacwest Bancorp | -59 | 0.86 | -155 | 0.70 |
| Rockville Financial, Inc. | -77 | 0.30 | -258 | 0.35 |
| Seacoast Banking Corporation Of Florida | -267 | 0.43 | -246 | 0.67 |
| Southside Bancshares, Incorporated | -4 | 0.93 | -145 | 0.39 |
| Sun Bancorp, Inc | -241 | 0.67 | -589 | 0.28 |
| Svb Financial Group | 39 | 0.81 | 40 | 0.46 |
| Synovus Financial Corp. | -289 | 0.74 | -140 | 0.71 |
| Umb Financial Corporation | -342 | 0.30 | -470 | 0.33 |
| United Bankshares, Inc. | 737 | 0.39 | 53 | 0.87 |
| Washington Trust Bancorp, Inc. | 330 | 0.39 | 373 | 0.57 |
| Wesbanco, Inc. | 187 | 0.33 | 145 | 0.41 |
| Wintrust Financial Corporation | 9 | 0.93 | 544 | 0.19 |

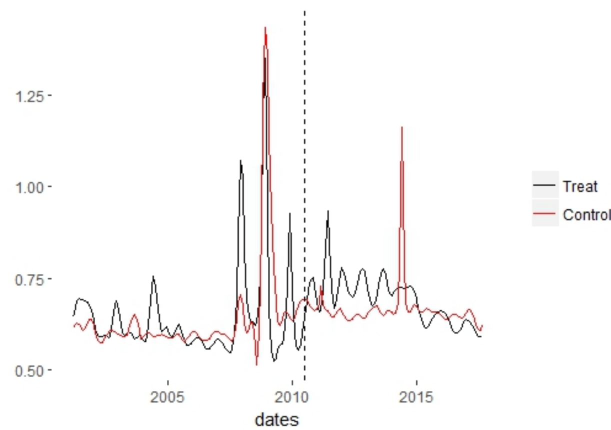
Table 2.24: Estimates of gaps in AFS securities as a percentage of Total Assets. Big Banks. Calculated as observed minus predicted. Bank by bank calculation

| Banks | Pre-Final Rule Avg Gap | Pre-Final Rule P-value | Post-Final Rule Avg Gap | Post-Final Rule P-value |
|---|---|---|---|---|
| Bank Of America Corporation | -0.0216 | 0.82 | -0.0627 | 0.67 |
| Citigroup Inc. | 0.0227 | 0.74 | 0.0195 | 0.81 |
| Citizens Financial Group Inc | -0.1443 | 0.13 | -0.2781 ** | 0.01 |
| Comerica Incorporated | 0.0067 | 0.92 | 0.0084 | 0.89 |
| Fifth Third Bancorp | -0.0711 | 0.71 | -0.0639 | 0.79 |
| Jpmorgan Chase & Co. | -0.0242 | 0.74 | -0.0701 | 0.43 |
| Keycorp | 0.0114 | 0.76 | 0.1157 * | 0.07 |
| M&T Bank Corporation | -0.0445 | 0.29 | 0.0099 | 0.89 |
| Northern Trust Corporation | -0.0581 | 0.74 | -0.0706 | 0.79 |
| Pnc Financial Services Group, Inc., The | 0.0026 | 0.97 | 0.0299 | 0.81 |
| State Street Corporation | -0.0325 | 0.94 | 0.1643 | 0.71 |
| Suntrust Banks, Inc. | -0.0829 | 0.15 | -0.1213 * | 0.07 |
| U.S. Bancorp | -0.0423 | 0.29 | -0.0217 | 0.74 |
| Wells Fargo & Company | 0.0564 | 0.63 | 0.0295 | 0.89 |
| Zions Bancorporation | -0.0531 | 0.29 | -0.0253 | 0.33 |

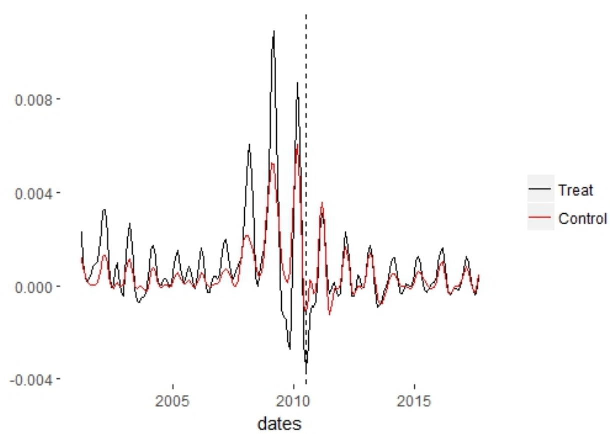
Table 2.25: Estimates of gaps in AFS securities as a percentage of Total Assets. Small Banks. Calculated as observed minus predicted. Bank by bank calculation

| Banks | Pre-Final Rule Avg Gap | Pre-Final Rule P-value | Post-Final Rule Avg Gap | Post-Final Rule P-value |
|---|---|---|---|---|
| Auburn National Bancorporation, Inc. | -0.0875 | 0.39 | -0.0895 | 0.51 |
| Banner Corporation | 0.008 | 0.60 | 0.0191 | 0.89 |
| Bok Financial Corporation | -0.0255 | 0.81 | -0.0419 | 0.79 |
| Boston Private Financial Holdings, Inc. | -0.0471 | 0.51 | -0.0141 | 0.82 |
| Camden National Corporation | 0.0659 | 0.51 | 0.0177 | 0.81 |
| Cathay General Bancorp | -0.1755 | 0.28 | -0.1382 | 0.51 |
| Chemung Financial Corp | -0.1518 | 0.22 | -0.2469 | 0.11 |
| Colony Bankcorp, Inc. | 0.0851 * | 0.07 | 0.1436 ** | 0.01 |
| Community Bank System, Inc. | 0.0106 | 0.75 | 0.24 | 0.11 |
| F.N.B. Corporation | 0.0878 | 0.51 | 0.0743 | 0.71 |
| First Financial Bancorp | 0.0416 | 0.83 | -0.0182 | 0.97 |
| First Financial Bankshares, Inc. | -0.0664 | 0.46 | 0.2006 | 0.14 |
| First Horizon National Corporation | 0.0102 | 0.82 | 0.0323 | 0.57 |
| First United Corporation | -0.0026 | 0.86 | -0.0468 | 0.74 |
| Heritage Financial Corporation | 0.0097 | 0.58 | 0.2152 *** | 0.00 |
| Investors Bancorp, Mhc | -0.0352 | 0.99 | -0.0705 | 0.90 |
| Mainsource Financial Group, Inc. | 0.0182 | 0.86 | -0.0178 | 0.92 |
| Old National Bancorp | -0.0839 | 0.29 | -0.0456 | 0.74 |
| Old Second Bancorp, Inc. | 0.0854 | 0.28 | 0.1244 | 0.24 |
| Pacwest Bancorp | 0.0776 | 0.38 | -0.0855 | 0.43 |
| Rockville Financial, Inc. | -0.0461 | 0.57 | 0.0478 | 0.69 |
| Seacoast Banking Corporation Of Florida | -0.0379 | 0.71 | 0.0317 | 0.89 |
| Southside Bancshares, Incorporated | 0.0946 | 0.39 | -0.0614 | 0.60 |
| Sun Bancorp, Inc | -0.122 | 0.46 | -0.0538 | 0.79 |
| Svb Financial Group | 0.1771 | 0.29 | 0.035 | 0.79 |
| Synovus Financial Corp. | 0.0144 | 0.46 | 0.0117 | 0.69 |
| Trico Bancshares | -0.0316 | 0.57 | -0.1121 | 0.38 |
| Umb Financial Corporation | -0.0165 | 0.92 | -0.0466 | 0.86 |
| United Bankshares, Inc. | -0.0348 | 0.65 | -0.0269 | 0.79 |
| Washington Trust Bancorp, Inc. | -0.0912 | 0.43 | -0.1532 | 0.35 |
| Wesbanco, Inc. | -0.0367 | 0.78 | -0.0003 | 0.92 |
| Wintrust Financial Corporation | -0.1569 | 0.22 | -0.2276 | 0.21 |


(a) Profitability

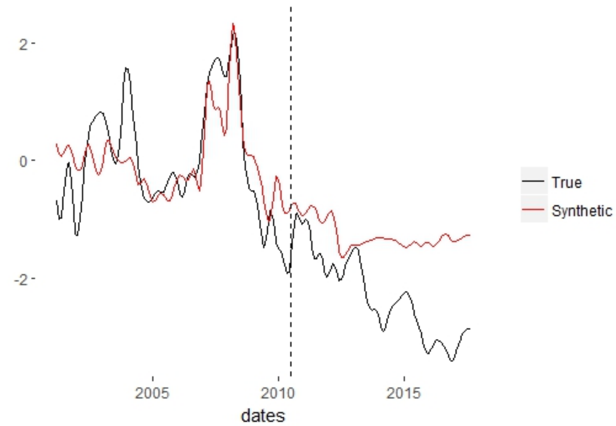



(b) Efficiency

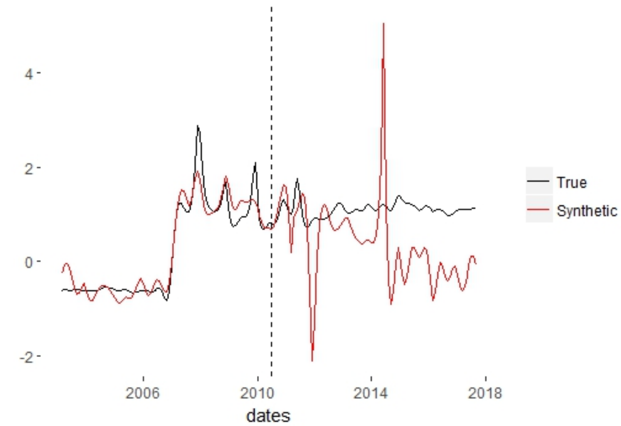



(c) Asset Quality

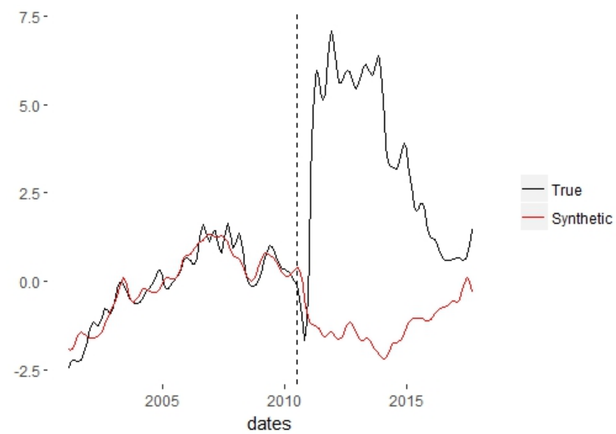
Figure 2.1: Covariates. Total Asset weighted average




(a) Trading Assets




(b) IB advisory, brokerage, and underwriting



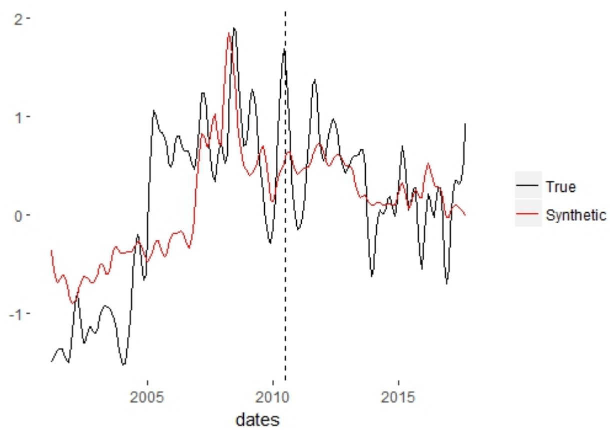

(c) Total GNA of derivative contracts held for purpose other than trading


(d) Total GNA of derivative contracts held for trading

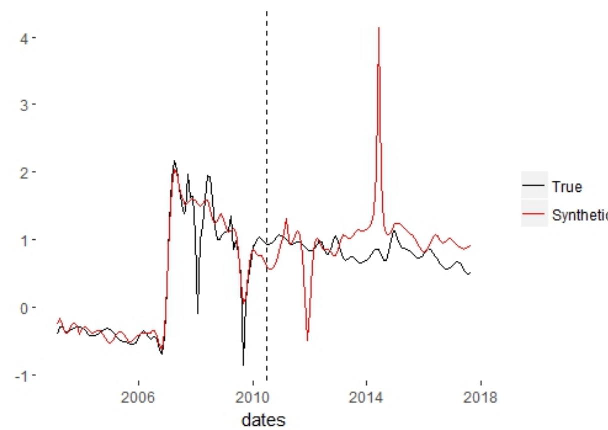

(e) Z-score


(f) Available for sale securities

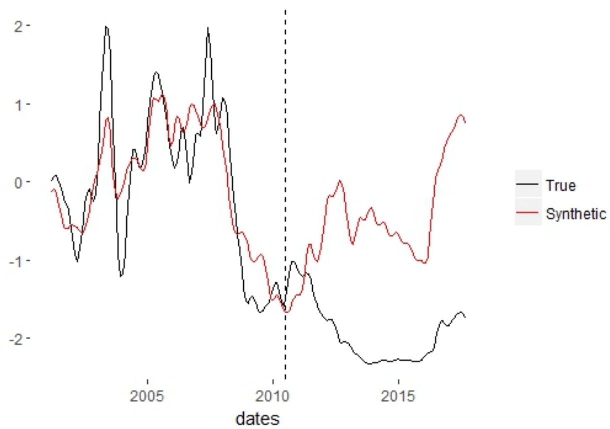
Figure 2.2: Synthetic control vs true values. Representative big bank



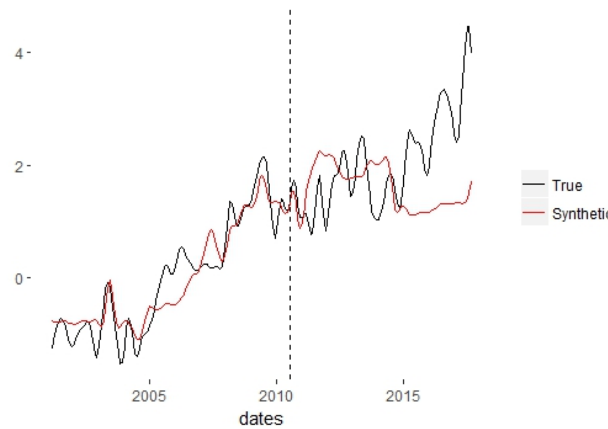

(a) Trading Assets



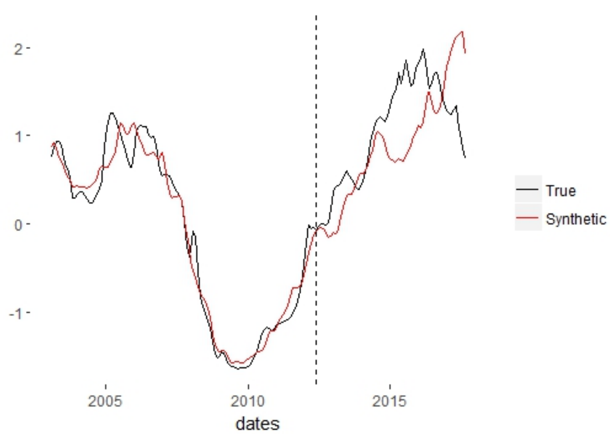

(b) IB advisory, brokerage, and underwriting



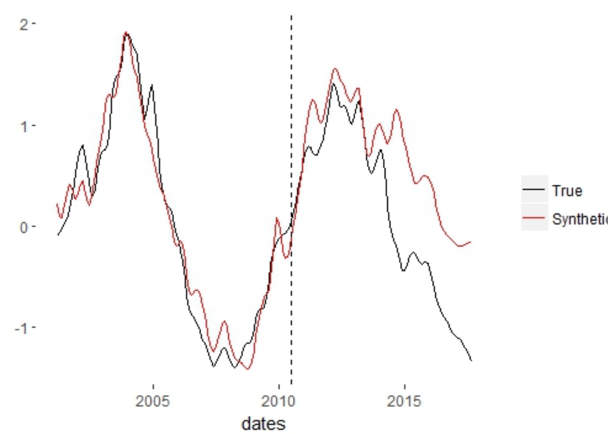

(c) Total GNA of derivative contracts held for purpose other than trading




(d) Total GNA of derivative contracts held for trading




(e) Z-score




(f) Available for sale securities

Figure 2.3: Synthetic control vs true values. Representative small bank

# Bibliography


Abadie, A., Diamond, A., and Hainmueller, J. (2007). Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. Working Paper 12831, National Bureau of Economic Research.

Abadie, A., Diamond, A., and Hainmueller, J. (2015). Comparative politics and the synthetic control method. *American Journal of Political Science*, 59(2):495–510.

Abadie, A. and Gardeazabal, J. (2003). The economic costs of conflict: A case study of the Basque country. *American Economic Review*, 93:113–132.

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets*, 5(1):31 – 56.

Avery, C. and Zemsky, P. (1998). Multidimensional Uncertainty and Herd Behavior in Financial Markets. *The American Economic Review*, 88(4):724–748.

Baily, M. N., Klein, A., and Schardin, J. (2017). The Impact of the Dodd-Frank Act on Financial Stability and Economic Growth. *RSF: The Russell Sage Foundation Journal of the Social Sciences*, 3(1):20–47.

Banerjee, A. V. (1992). A Simple Model of Herd Behavior. *The Quarterly Journal of Economics*, 107(3):797–817.

Bao, J., O'Hara, M., and Zhou, X. (2016). The Volcker Rule and Market-Making in Times of Stress. Finance and Economics Discussion Series 2016-102, Board of Governors of the Federal Reserve System (US).

Barndorff-Nielsen, O. E. (2004). Measuring the impact of jumps in multivariate price processes using bipower covariation.

Bessembinder, H., Jacobsen, S., Maxwell, W., and Venkataraman, K. (2018). Capital Commitment and Illiquidity in Corporate Bonds. *The Journal of Finance*, 73(4):1615–1661.

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1992). A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. *Journal of Political Economy*, 100(5):992–1026.

Bikhchandani, S. and Sharma, S. (2000). Herd behavior in financial markets : A review. *IMF Working Papers*.

Breiman, L., Friedman, J., Olshen, R., and Stone, C. (1984). *Classification and Regression Trees*. Wadsworth and Brooks, Monterey, CA.

Cambria, E. and White, B. (2014). Jumping NLP Curves: A Review of Natural Language Processing Research [Review Article]. *IEEE Computational Intelligence Magazine*, 9(2):48–57.

Campbell, J. Y. (2006). Household Finance. *The Journal of Finance*, 61(4):1553–1604.

Card, D. and Krueger, A. B. (1994). Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania. *The American Economic Review*, 84(4):772–793.

Chatterjee, R. (2011). Dictionaries Fail: The Volcker Rule's Reliance on Definitions Renders it Ineffective and a New Solution is Needed to Adequately Regulate Proprietary Trading. *Brigham Young University International Law & Management Review*, 8(1):33–62.

Chow, J. and Surti, J. (2011). Making banks safer : Can volcker and vickers do it? *IMF Working Papers*.

Chung, S., Keppo, J., and Yuan, X. (2019). The Impact of Volcker Rule on Bank Profits and Default Probabilities. SSRN Scholarly Paper ID 2167773, Social Science Research Network, Rochester, NY.

Cipriani, M. and Guarino, A. (2014). Estimating a Structural Model of Herd Behavior in Financial Markets. *The American Economic Review*, 104(1):224–251.

Cramer, J. S. (1999). Predictive performance of the binary logit model in unbalanced samples. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 48(1):85–94.

Dick-Nielsen, J. and Rossi, M. (2016). The cost of immediacy for corporate bonds. SSRN Scholarly Paper ID 2139403, Social Science Research Network.

Doudchenko, N. and Imbens, G. W. (2016). Balancing, Regression, Difference-In-Differences and Synthetic Control Methods: A Synthesis. *arXiv:1610.07748 [stat]*. arXiv: 1610.07748.

Drehmann, M., Oechssler, J., and Roider, A. (2005). Herding and Contrarian Behavior in Financial Markets: An Internet Experiment. *American Economic Review*, 95(5):1403–1426.

Duffie, D. (2012). Market Making Under the Proposed Volcker Rule. SSRN Scholarly Paper ID 1990472, Social Science Research Network, Rochester, NY.

Estabrooks, A., Jo, T., and Japkowicz, N. (2004). A multiple resampling method for learning from imbalanced data sets. *Computational Intelligence*, 20(1):18–36.

Flach, P. and Ferri, C. (2011). A coherent interpretation of auc as a measure of aggregated classification performance. In *In Proceedings of the 28th International Conference on Machine Learning, ICML2011.*

Fluss, R., Faraggi, D., and Reiser, B. (2005). Estimation of the youden index and its associated cutoff point. *Biometrical Journal*, 47(4):458–472.

FSOC (2011). Study recommendations on prohibitions on proprietary trading certain relationships with hedge funds private equity funds. *Proprietary Trading and Investment Restrictions Under the Volcker Rule*, pages 103–169.

Geman, S., Bienenstock, E., and Doursat, R. (1992). Neural networks and the bias/variance dilemma. *Neural Computation*, 4:1–58.

Getis, A. and Aldstadt, J. (2004). Constructing the Spatial Weights Matrix Using a Local Statistic. *Geographical Analysis*, 36(2):90–104.

Glosten, L. R. and Milgrom, P. R. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14(1):71 – 100.

Go, A., Bhayani, R., and Huang, L. (2009). Twitter sentiment classification using distant supervision. *Processing*, 150.

Hirshleifer, D. A. (2001). Investor Psychology and Asset Pricing. SSRN Scholarly Paper ID 1288971, Social Science Research Network, Rochester, NY.

Ho Lee, I. (1998). Market crashes and informational avalanches. *Review of Economic Studies*, 65:741–59.

in Lee, S., Lee, H., Abbeel, P., and Ng, A. Y. (2006). Efficient l1 regularized logistic regression. In *Proceedings of the 21st National Conference on Artificial Intelligence*.

J. Bloomfield, R., O'Hara, M., and Saar, G. (2007). How noise trading affects markets: An experimental analysis. *SSRN Electronic Journal*.

John, V. and Vechtomova, O. (2017). Sentiment analysis on financial news headlines using training dataset augmentation. *CoRR*, abs/1707.09448.

Keppo, J. and Korte, J. (2016). Risk Targeting and Policy Illusions - Evidence from the Announcement of the Volcker Rule. SSRN Scholarly Paper ID 2466477, Social Science Research Network, Rochester, NY.

Krawiec, K. D. and Liu, G. (2015). The Volcker Rule: a brief political history. *Capital Markets Law Journal*, 10(4):507–522.

Lakonishok, J., Shleifer, A., and Vishny, R. W. (1992). The impact of institutional trading on stock prices. *Journal of Financial Economics*, 32(1):23 – 43.

Loughran, T. and McDonald, B. (2011). When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *The Journal of Finance*, 66(1):35–65.

Lu, K. (2017). Growth in mobile news use driven by older adults. *Pew Research Center*.

Macey, J. R. (2000). The business of banking: Before and after Gramm-Leach-Bliley. *Journal of Corporate Law*, 25:691–722.

Madura, J. and Premti, A. (2014). Impact of the volcker rule on bank valuations and risk. *Banking and Finance Review*, 6(12).

McFedries, P. (2009). Information at your fingertips. *IEEE Spectrum*.

Mill, J. S. (2011). *A System of Logic, Ratiocinative and Inductive: Being a Connected View of the Principles of Evidence, and the Methods of Scientific Investigation*, volume 1 of *Cambridge Library Collection - Philosophy*. Cambridge University Press.

Oh, C. and Sheng, O. (2011). Investigating predictive power of stock micro blog sentiment in forecasting future stock price directional movement. In *ICIS*.

Park, A. and Sabourian, H. (2011). Herding and Contrarian Behavior in Financial Markets. *Econometrica*, 79(4):973–1026.

Rinkler, T. (2019). Package 'sentimentr'. *CRAN*.

Sun, A., Lachanski, M., and J. Fabozzi, F. (2016). Trade the tweet: Social media text mining and sparse matrix factorization for stock market prediction. *International Review of Financial Analysis*, 48.

Thakor, A. V. and E., S. J. (2012). The economic consequences of the volcker rule. Technical report, US Chamber of Commerce's Center for Capital Market Competitiveness.

Trebbi, F. and Xiao, K. (2015). Regulation and Market Liquidity. NBER Working Papers 21739, National Bureau of Economic Research, Inc.

Tsuruoka, Y., Tsujii, J., and Ananiadou, S. (2009). Stochastic gradient descent training for l1-regularized log-linear models with cumulative penalty. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 1 - Volume 1*, ACL '09, pages 477–485, Stroudsburg, PA, USA. Association for Computational Linguistics.

Wermers, R. (1999). Mutual fund herding and the impact on stock prices. *The Journal of Finance*, 54(2):581–622.

Whitehead, C. K. (2011). The Volcker Rule and evolving financial markets. *Harvard Business Law Review*.

Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer*, 3(1):32–35.

Zhang, T. and Oles, F. J. (2000). Text categorization based on regularized linear classification methods. *Information Retrieval*, 4:5–31.

Zhu, J. and Hastie, T. (2005). Kernel logistic regression and the import vector machine. *Journal of Computational and Graphical Statistics*, 14(1):185–205.

Zou, H. and Hastie, T. (2005). Regularization and Variable Selection via the Elastic Net. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 67(2):301–320.

OCC et al. (2018). Proposed revisions to prohibitions and restrictions on proprietary trading and certain interests in, and relationships with, hedge funds and private equity funds. *83 FR 33432*, pages 33432–33605.