Essays on Information, Markets, and Mandated Disclosures

By
Mehrzad Azmi Shabestari

A dissertation submitted to the

Graduate School-Newark

Rutgers, The State University of New Jersey

in partial fulfillment of requirements

for the degree of

Doctor of Philosophy

Graduate Program in Management

written under the direction of

Professor Bharat Sarath

and approved by

____

Dr. Bharat Sarath

____

Dr. Dan Palmon

____

Dr. Kevin Moffitt

____

Dr. John Shon

Newark, New Jersey
January 2017

ABSTEACT OF THE DISSERTATION

Essays on Information, Markets, and Mandated Disclosures

By Mehrzad Azmi Shabestari

Dissertation Director:
Professor Bharat Sarath

This dissertation consists of two essays. The first essay focuses on the ability of mandated SEC disclosures and voluntary corporate restrictions to restrain insiders from trading based on material, non-public information. We look at the combined effects of required Form-4 disclosures and the single most common corporate policy which restricts trading by insiders at all times except a short window starting three to 12 days after the earning announcement date which is called the white window. Results indicate that there is a strong selection bias by insider deciding which white window to trade in. Because of this selection bias there is not a significant difference between the average cumulative abnormal return of the black and white periods, but insiders gain much more cumulative abnormal return comparing the returns of the black period with the white periods that they do not trade in. We also compare the average cumulative abnormal return for the trades which are filed under a 10b5-1 plan versus non-plan trades. We find that in the black period, and in a short period right before earnings announcement date, the trades which are filed under a10b5-1 plan generate more cumulative abnormal return on average.

The second essay studies disclosures by the banking industry about potential business risk factors leading into the financial crisis of 2007- 08. Such disclosures are mandated in

Item 1A of the annual 10-K form filed with the SEC. The objective of this research is to evaluate the informativeness and timeliness of one specific component of annual reports (Form 10-K), namely Item 1A. We qualitatively examine this issue by checking the underlying tone of these disclosures. The results indicate that the tone of Item 1A has became much more negative from 2008 to 2009. We have also created a dictionary of financial crisis related words using LDA which is a generative statistical model to check the timing of appearance of those words in Item 1A. Overall, the results support the argument that the banking industry has failed to predict the financial crisis, and they have started noticing and disclosing risks related to the crisis only after it happened.

Acknowledgements

I would like to express my deepest gratitude to the members of my dissertation committee, Dr. Bharat Sarath, Dr. Dan Palmon, Dr. Kevin Moffitt, and Dr. John Shon.

Above all, I would like to thank my advisor Dr. Sarath for his guidance, encouragement, and continuous optimism. You have been such a great mentor and supporter throughout this process. You have made me not only a better researcher, but a better person. And I am beyond thankful for this opportunity.

Special thanks to Dr. Moffitt for his mentorship, patience, and continuous support.

Last but not least, I would like to thank my parents for their love and support. I could not have accomplished this without you.

Dedication

To my parents

Table of Contents

List of Tables

List of Figures

Chapter 1: Patterns of Insider Trading

1.1 Introduction

Insider trading is a term that can be subject to different definitions. It can be legal or illegal depending on the time that the trade is executed, and whether the trade is based on confidential information. Insider trading takes place legally every day, when corporate insiders – officers, directors or employees – buy or sell stock in their own companies within the confines of company policy and the regulations governing this trading.[1] Illegal insider trading happens when the insider of the company takes advantage of private information about important events within the company to make more profit or avoid losses on the stock market (Newkirk and Robertson 1998).

This study focuses on the ability of mandated SEC disclosures and voluntary corporate restrictions to restrain insiders from trading based on material, non-public information. We look at the combined effects of required Form-4 disclosures and the single most common corporate policy which restricts trading by insiders at all times except a short window starting three to 12 days after the earning announcement date which is called the white window. Results indicate that there is a strong selection bias by insiders deciding which white window to trade in. Because of this selection bias, there is not a significant difference between the average cumulative abnormal return of the black and white periods, but insiders gain much more cumulative abnormal return comparing the black period with the white periods that they do not trade in. We also compare the average cumulative abnormal return for the trades which are filed under a 10b5-1 plan versus

---

[1] https://www.sec.gov/news/speech/speecharchive/1998/spch221.htm

non-plan trades. We find that in the black period, and in a short period right before earnings announcement date, the trades which are filed under a10b5-1 plan generate more cumulative abnormal return on average.

The rest of this dissertation is laid out as follows. In section 2, we discuss the insider trading literature review and regulatory background to insider trading. Then, we develop the hypotheses. In section 3, we describe the data and sample selection. Research design and empirical tests are presented in section4. The results are provided in section 5. In section 6 we look at the companies that have internal insider trading policies while the conclusion is provided in section 7.

1.2 Literature Review and Hypotheses Development

1.2.1 Insider Trading Background

Insider trading laws are a matter of contention both from the point of economic theory and from the perspective of legal enforcement. In a series of articles, Manne (see for example Manne [1973]) argued that curbs on insider trading interfered with the efficiency of markets. In contrast other scholars (see for example Fishman and Hagerty [1992]) suggest that the fear of insiders in the stock market results in reduced information in prices. The basic tension may be summarized as follows: Does insider trading increase the flow of information rendering it more efficient, or does it act as a disincentive for others to gather information or to trade reducing the information reflected in prices?

In addition to this mainly theoretical dispute, there is a practical consideration arising from the fact that stock based compensation schemes have become popular for top executives. For these schemes to be effective, insiders have to be permitted to sell their shares and convert (at least part of) their holdings into cash. Given these competing economic pressures, the SEC has mostly followed a policy of "disclose or abstain" where insiders are allowed to trade as long as these trades are surrounded by sufficient disclosures that "mitigate the insider's informational advantage". The goal this chapter is to study how well this approach is working in practice and whether insiders still benefit from an informational advantage based on the performance of their actual trades.

One reason why SEC's rules and other curbs on insider trading (such as corporate policies) may have unpredictable outcomes is related to the legal uncertainty in the interpretation of insider trading statutes. In recent years, the New York attorney general's office has targeted insider trading with considerable energy. After a long period where the government won almost every insider trading case that they chose to pursue, the tide turned abruptly in 2013 with the defendants being acquitted in a number of high profile cases.2 Given the complexity in deciding what actually constitutes illegal insider trading, the majority of SEC actions pertain to cases involving buying shares ahead of a merger. In a parallel fashion, investor suits are usually part of a class-action where the accusation of selling ahead of bad news is used as a technique for proving that managers had knowledge of bad news that they did not disclose in a timely fashion. In other words, insider trading accusations are not important in their own right but rather as a step in

---

[2] The recent acquittal in 2014 of two indicted traders, Newman and Chaisson, was summed up by Forbes as follows:
One of the requirements for insider trading requires the person providing a confidential tip – known as a tipper – to receive a benefit for divulging confidential information. During oral arguments in April, the Second Circuit Court of Appeals cast doubt on the government's prosecution of downstream tippees – recipients of confidential information who are more than one step removed from the tippers conveying the illicit information. The three-judge panel overseeing a case involving Todd Newman and Anthony Chiasson hinted that prosecutors would have to prove that a tippee knew that the tipper received a benefit for passing on confidential information. Up to that point, the government had successfully prosecuted tippees without having to prove that they were cognizant of such benefits.

demonstrating that insiders held information that they failed to disclose to others. Incidentally, this would also violate insider trading statutes of 1934 that forbid trading based on any material information that is not available to the public.

In the absence of actual lawsuits or direct SEC action, the main restrictions on insider trading arise from indirect pressure on corporate policies and brokers. The Insider Trading and Securities Fraud Enforcement Act of 1988 ("ITSFEA"), was designed primarily to prevent, deter, and prosecute insider trading. Section 15 (f) of the Securities Exchange Act of 1934, created pursuant to promulgation of ITSFEA,[3] requires broker-dealers to maintain procedures to prevent the misuse of material non-public information. Passage of this Act made firms responsible for violations of insider trading by their officers. Consequently, many firms have adopted voluntary insider trading regulations to prevent their insiders to trade based on inside information. Bettis, Coles and Lemmon (2000) use a survey approach to compare the profitability of insider trading between firms which have voluntary corporate restrictions for insider trading and the firms which do not have any restrictions. The "acceptable window for trade" (hereafter, "white window") was deemed to be 3-12 trading days after the earnings announcement when insiders are assumed to have the least informational advantage. Bettis et al. (2000) shows that a number of firms implemented such policies and that trading restrictions can be beneficial because it leads to smaller bid- ask spreads and greater liquidity in the windows that the insiders are not allowed to trade in (hereafter, "black window").

[3] See (https://www.sec.gov/divisions/marketreg/brokerdealerpolicies.pdf)

The approach in this paper is based on this same idea of a "low information asymmetry" white window that spans days 3-12 after an earnings announcement. Prior literate shows the significant effect of earnings announcement on price and volume (Beaver, 1968), and significant changes in bid-ask spreads (Lee et al., 1993). Kim and Verrecchia (1994) suggests that bid- ask spread temporarily widen around earnings announcements. They argue that market makers are likely to increase spreads in anticipation of an earnings announcement, to protect themselves against insiders trading on the information before it is disclosed publicly. Krinsky and Lee (1996) results also support the increase in the information asymmetry before earnings releases.

As documented in Beaver (1968) or Patell and Wolfson (1981,1984) ,the stock price is more volatile in the period before the earnings announcement when there is more uncertainty. Beaver (1968) suggests that the magnitude of the price change (without respect to sign) should be larger immediately before the earnings announcement becomes public because investors take different positions based on their private beliefs regarding the impending news. Further , when the earnings news is released to the public, uncertainty is eliminated and we observe a convergence in beliefs and less volatility in the stock price. The rationale behind the white, black, and the blackest windows is based on this regular pattern of price uncertainty. The White window, only starts three days after the earnings announcement by which time market prices should fully reflect the information disclosed. Therefore, this is the window with lowest information asymmetry and insiders are allowed to trade freely in this window because they are unlikely to enjoy any advantage. In the UK, the 1985 Companies' Act specifies that directors are prohibited from dealing in the securities of their own companies for a period of two months prior to the announcement of year-end or half-year results, and at other times prior to the announcement of price-sensitive information. The British statute is more comprehensive and would typically preclude any trading in the “Black Window” in the two weeks preceding an earnings announcement.

The rest of the period leading up to the next earnings announcement is split into a “moderate information asymmetry” black window and a “high information asymmetry” blackest window. We examine a number of price and trade based metrics to see if there are systematic differences across the black and white windows. The focus of this analysis is threefold:

1. Whether trades executed in the white windows display any evidence of informational advantages for insiders;
2. Whether trades executed in the black windows display evidence of informational advantages for insiders;
3. Is the evidence for an informational advantage stronger in the black windows than the white windows?

Another important piece of regulation aimed at striking a balance between permitting insiders to diversify their holdings and restraining them from exploiting private information was the creation of registered 10b5-1 plans. Under these plans, insiders file a long-term plan for liquidating their holdings on a periodic basis. 10b5-1 plans typically specify the number of shares to be sold as well as floor prices for these sales. The premise underlying these plans is that by fixing a date in advance, it will be a matter of random chance whether insiders hold material private information at the time of the sale that could take place many years down the road. However, as argued in recent papers such as

Jagolinzer (2009, 2013) and Shon and Veliotis (2013), 10b5-1 plans can actually turn into vehicles for riskless insider trading.

Jagolinzer (2009) argues that both the initiation and termination of 10b5-1 are strategically timed to take advantage of private information. For example, a plan may be filed in advance of adverse information and then used to sell stock prior to the information becoming public. Later on, the plan could be wound up without consequences if the private information is positive. A less extreme example would simply be the cancellation of trade if the price is likely to be higher in the near future. As a matter of law, it is impossible to prosecute a party for insider trading if no trades have been executed. So a 10b5-1 plan gives a safe "option" for exploiting private information – sell under the plan if the private information suggests that the stock price is likely to dip in the future but refrain from selling if the stock price is likely to rise in the near future. This is exactly like an option where only profitable trades are exercised. Following on the logic of Jagolinzer (2009), I examine a fourth issue:

4. Are 10b5-1 plans more likely to be scheduled in black windows rather than white windows?

The thesis in Jagolinzer (2009, 2013) is that 10b5-1 plans are not meeting the goal of preventing insiders from exploiting private information. One simple method for achieving this is to schedule 10b5-1 plans in the blackest window where the private information is highest and the option to trade (or refrain) has its highest value.

Another possibility with regard to insider trading, advances in Shon and Veliotis (2013) is that firms with 10b5-1 plans are more likely to manage earnings to maximize the option value associated with these plans. Specifically, if a firm is able to raise its short-term short price through earnings management, we should expect to see a pattern of buys before the earnings announcements and sales after. On the other hand, if the firm is not managing earnings, we should expect to see buys and sells based on private information before the earnings announcement but no trading after the earnings announcements. In other words, the relative volume of trade in the blackest window (prior to earnings announcements) compared with that in the white window (just after the earnings announcement) provides evidence on the proportion of profits on insider trading accrued through earnings management relative to that collected through exploiting private information.

The first point, noted already, is that there is significant trading in both the white and black windows. This suggests that Section 15(f) making corporations responsible for insider trading is not a significant deterrent in terms of restricting trading to the white window identified by Bettis (2000). The second point is that trading under 10b5-1 is much more likely in the blackest window (immediately preceding the earnings announcement) confirming the point advanced by Jagolinzer (2009, 2013) that these plans are serving as strategic vehicles for profitable insider trading. The strategic value of a 10b5-1 plan derives from the following strategy: if no private information (or adverse private information) is present, refrain from trade; otherwise, trade at a profit. As the probability that private information will be present is highest in the period before an earnings announcement, the value of this strategy is maximized by scheduling the plan in the period immediately preceding an earnings announcement.

In terms of market adjusted return earned by insiders, we find that trading is equally profitable in the white and black windows. A superficial interpretation might be that the whole notion of periods of low and high information asymmetry is suspect. However, a closer examination suggests a more interesting possibility. We examine the profits from randomized selling in all white windows (as opposed to only those where trading took place) and compare it with white window profits on actual trades. If indeed there was no private information available to insiders in the white window, a randomized trading strategy should earn the same returns as any other strategy. Instead, we find that the return on actual trades is much higher than that of randomized trades. That is, on the whole, the market price is efficient inside white windows (as conjectured in many earlier studies), but insiders select to trade only in white windows where the price fails to fully reflect the private information.

The findings suggest that the conventional viewpoint of windows of low and high information asymmetry is correct, but that insiders profit by trading in both these windows. The evidence also suggests that trading immediately before an earnings announcement is common, but insiders often choose to protect themselves by filing 10b5-1 plans in this period. An obvious question in this context is whether the insider informational advantage would be preserved even after the earnings announcement where there would be much less risk of running afoul of insider trading statutes. To answer this, we further examine what would happen if a trade made in a black window is postponed to the next available white window. In this scenario, we find that the trading profits would be eliminated, suggesting that there is some merit to statutes as those in the UK that preclude trading in the black windows.

There are two further issues that we will examine in this dissertation. The first is to see if the pattern of buying in the black window and selling in the next white window (or vice versa) leads to significant trading profits and whether this can be linked to strategies of meeting or beating expectations. A second issue is to examine differences across white and black windows using the partitioning of insider trading into routine trades and informative trades developed in Cohen, Malloy, and Pomorski (2012). We expect that the routine trades are more prevalent in white windows whereas informative trades are more frequent in black windows.

1.2.2 Regulatory Background

According to Rule 10b-5 of the Securities and Exchange Act of 1934, any trade based on material and non-public information is prohibited in the USA. Rule 10b5-1 was adopted in August 2000 not only to deter insiders from trading based on private information, but also to protect insiders' preplanned, non-information-based trades from litigation. Passage of SOX also had a significant influence on insider trading regulations. For instance, in the post-SOX period, insiders are required to report their trade to the SEC by filing a Form 4 within two business days of the trade, whereas before that they had to file their Form 4 within the first ten days of the next month which used to give them a longer time to report the trade. Besides these regulations that are enforced by the SEC, there are some company-level regulations of insider trading that are enforced by the companies themselves. Based on Bettis et al. (2000) the single most common corporate policy disallows trading by insiders at all times except during a trading window that is open during the period extending from three to 12 trading days after the quarterly earnings announcement.

1.2.3 **Hypotheses Development**

The focus of this study is on the blackout window corporate policy employed by some companies to prevent insider trading in periods of high information asymmetry. Bettis et al. (2000) indicate that blackout period policies have successfully prevented black period insider trading among surveyed firms. Since some firms in our sample have employed such a corporate policy, we expect to observe that trades which are executed during the black period generate larger CAR.

We conclude the first hypothesis as:

Hypothesis 1: Average cumulative abnormal return in the black and the white periods are the same.

The rationale behind this corporate policy is that there is less information asymmetry in the white periods, and insiders do not have any information advantages in the white periods over other investors. That is the reason trading is allowed for insiders in the white window. Based on this explanation, insiders should have no information advantage in white windows, and hence:

Hypothesis 2: White window returns to insiders should be insignificant.

However, if white window returns are significant, the question arises as to whether this is true for white windows in general or only for the windows where trades place. That is, insiders select to trade only in those white windows where the market price fails to fully reflect their private information.

Hypothesis 2-1: White window return in the traded period is the same as average of other white window returns.

Insiders risk trading in black periods to avoid losing an information advantage. This leads me to the third hypothesis that insiders will lose their information advantage if they play safe and wait to execute their trades in the white period.

Hypothesis 3: Trades that are executed in a period immediately prior to an earnings announcement generate the same cumulative abnormal return if delayed by ten days.

The presumption underlying the idea of a "proper" window for insider trading is that information advantages become strong outside that window.

Jagolinzer (2009) finds that insiders execute sales under pre-registered 10b5-1 plans before bad news, generating abnormal forward-looking returns. On the other hand Sen (2008) finds no significant differences in stock price performance following plan and non-plan sales. This leads me to the last hypothesis to check the difference between 10b5-1 trades and non-plan trades.

Hypothesis 4: Trades that are executed under a 10b5-1 plan generate the same average cumulative abnormal return in black/ white/ blackest periods comparing with the trades that are not executed under a 10b5-1 plan.

1.3 Data

1.3.1 Insider Trading Data

The insider trading data for this research consists of two tables collected from the Thomson Reuters TFN Insiders Data which covers the period starting from January 1, 1986. Table I contains non-derivative transactions and holdings information filed on Forms 3, 4, and 5. Table II contains derivative transactions and holdings information filed on Forms 3, 4, and 5. Officers, directors, and large shareholders who owns more than ten percent of a company's shares are required by Section 16(a) of the Securities and Exchange Act of 1934 to report any transaction to the SEC on forms 3, 4, and 5. According to SEC regulations, the initial filing should be reported on Form 3 within ten days of becoming an officer, director, or beneficial owner. Changes in ownership should be filed on a Form 4 within two business days after a trade. Any transactions that should have been reported on a Form 4 earlier or were eligible for deferred filing should be filed on a Form 5. In this paper we focus on form 4s which are reported under Table I, as those form 4s are the ones documenting insider trading. We run the analysis for two different time periods: 1998 to 2002 (pre-SOX), and 2003 to 2006 (post-SOX). The rationale behind analyzing the two time periods separately is that "SOX has significantly affected many corporate governance provisions, including regulations related to insider trading" (Brochet, 2010). SOX requires insiders to file a Form 4 within two business days after a trade, rather than ten days after the start of the following calendar month. We include data until 2006 to make sure the results are not affected by the financial crisis by any means. The sample consists of 5,221,619 transactions made by 197,922 insiders in 18,252 companies from 1998 to 2006. We only include the transactions which has an "H" or "R" indicator for the variable "Cleanse". "H" means all data cleansing updates were made with high confidence and "R" means record passed all cleansing checks for reasonableness, and this reduces the sample to 3,270,790 transactions. The sample includes only open market transactions (transaction codes P and S). We end up with 1,955,851 transactions, 783,256 pre-SOX period and 1,172,595 post-SOX. Overall the variables collected from TFN files are as following:

1. Company name and CUSIP

2. Person ID which is an identical id for each insider
3. Form Type (Form 4), and Transaction Code ("P" for open market purchase and "S" for open market sale)
4. Role code which is the insider's role or position within the company. Role code can be classified into groups of directors, committees, officers, beneficial owners, and others
5. Transaction date, transaction price, and number of shares exchanged in the transaction
6. Cleanse ("H" for all data with high cleansing confidence and "R" for records passed all cleansing checks for reasonableness)

The earnings announcement data is collected from Compustat Industrial Quarterly filings. We align each transaction between two earnings announcement dates. The former earnings announcement date is the one that we will use to place the transaction in the timeline. Variable "timediff" is defined to count the number of trading days between the transaction date and the related earnings announcement date. For the number of trading days to be counted correctly, we put together a dataset with all the stock market holidays and exclude those when computing "timediff". Values for the variable "timediff" can vary from zero to 59, since each quarter consists of 60 trading days. Zero timediff for a transaction shows that the transaction has occurred on the same day as the earnings announcement date. All transaction with timediff value between 3 and 12 are considered "white period" trades, all the others will be "black period" trades. Closing price data for all the companies is downloaded from CRSP daily files.

1.3.2 Sample Selection

Total number of transactions from Jan 1998 to Dec 2006	5,221,619
1) Eliminate transactions with cleanse indicator other than "H" or "R"	1,950,829
	3,270,790
2) Eliminate transactions with form type other than Form 4	168,557
	3,102,138
3) Eliminate transactions with trancode other than "P" or "S"	1,146,287
	1,955,851
4) Eliminate transactions with of shares less than one	1,201
	1,954,650
5) Eliminate transactions with no matching earnings announcement date	342,512
	1,612,138
6) Eliminate transactions that do not have CRSP data	85,923
	1,526,215
7) Eliminate transactions that do not have COMPUSTAT data	73,937
	1,452,278
Pre-SOX Period Sample Size	565,983
Post-SOX Period Sample Size	886,295

1.3.3 10b5-1 Data

J3 Services Group (J3SG) provides additional information on relevant trades. Unlike Thomson Reuters, J3SG distinguishes 10b5-1 and non-10b5-1 transactions. We only consider “open market sale” and “open market purchase” transactions filed under a form 4 for analysis. This database covers open market sales and purchases from 2004 to 2014 and includes a flag to distinguish between 10b5-1 and non-10b5-1 trades. Based on Table 9, there are 736,916 open market sales and purchases from 2004 to 2014, out of which 229,276 (31%) are filed under a 10b5-1 plan. We only include the 359,449 trades that can be matched with CRSP and COMPUSTAT, 168,908 of which are filed under a 10b5-1 plan. We use the same method of aligning the trades in the timeline of one quarter using the earnings announcement date data from COMPUSTAT. We identify each trade as it has been executed in the white or black period. Because the black window is much longer than the white window, we identify a shorter version of the black period based on timediff variable and we call it the “blackest” window. All trades which are executed in the last ten days before the earnings announcement date; which we believe the most information asymmetry is available; are flagged as “blackest” trades. Table 1.10 shows how the 10b5-1 plans are spread throughout the quarter based on the defined windows. On average 19.51% of 10b5-1 sells are executed in the white period, 80.49% in the black period, and 34.50% in the blackest period. Panel B of Table 1.10 reports that on average 17.57% of 10b5-1 purchases are filed in the white period, 82.43% in the black period, and 26.84% in the blackest period.

1.4 Research Design

1.4.1 Empirical Tests

According to Patell and Wolfson (1981,1984) and Beaver (1968), stock price is more volatile in the period before the earnings announcement when there is more uncertainty. And when the news is released to public, uncertainty is eliminated and we observe less volatility in the stock price. The rationale behind the white, black, and the blackest windows is based on this finding. Before analyzing insider trading, we examine the differences in returns between the white and the blackest window. We download the closing price for all the companies from CRSP, and the information for earnings announcement date is downloaded from Compustat quarterly filings. Aligning the earnings announcement data with the observations from CRSP, we identify the closing price for each company on each day in the quarter. We only keep the information for beginning and ending days of white period (3 and 12), and the beginning and the ending days of blackest period (50 and 59). Then, we calculate the return for the black and white periods to check the difference in return for that windows.

According to Bettis et al. (2000) the corporate policy of “black windows” has successfully prevented trading in the periods with presumably more information asymmetry between insiders and other traders. Their results indicate that insider trading is concentrated heavily during the windows in which trading is permitted, with insider trading activity in the blackout period at less than one-third of that during allowed periods. Figure 9 shows that on average 67.69% of all trades have been made in the period that trading is not permitted from 1998 to 2006. This concentrated trading in the black period when more information asymmetry is present should result in higher abnormal returns for insiders, based on definition. We check cumulative abnormal return for both of the pre-SOX and post-SOX time periods. The purpose of this test is to check whether or not the black period insider trades' cumulative abnormal return is similar to the white period. We use a standard event study methodology for this test, defining the event date as the day the insider has made the trade either sale or purchase. The event study is an approach to calculate the amount of abnormal return based on the event date for a period of time. Each security in the sample is regressed for a time series of daily returns against the yields from a market index using this equation:

$$R_t = \alpha + \beta R_{mt} + e_t$$

Where R_t denotes the return on the security for time period t, R_{mt} denotes the return on a market index for period t, and e_t represents a firm-specific return (Lintner, 1965; Sharpe, 1963, 1964).

The next test is designed to check whether all the white periods are the same regarding the ability of generating cumulative abnormal return or is there a selection bias in choosing which white period to trade in. We compare the CARs of the traded and non-traded white periods for each individual insider. We have already calculated the average CAR for all the white periods that an individual insider has traded in. To calculate the average CAR of other white periods that the insider did not trade in, we get the average of CAR for first, one day in the middle, and last days of the white period (third day, 8th day, and 12th day of a quarter). This gives the average CAR for the white periods that the insider decides not to trade in.

We also design another setting to check the differences between white and black period CARs. Because the white window is only ten days long, we take the last ten days of a black period right before the beginning of the white period and map those trades to the following white period. By this setting, we want to check how the CAR would have changed if the insiders would have waited to make their trades in the white period instead of making them in the black period.

1.4.2 10b5-1 Tests

Jagolinzer (2009) examines whether insiders strategically trade within the safe harbor provided by the U.S. Securities and Exchange Commission (SEC) Rule 10b5-1. Their data covers the period between October 2000 and December 2005, and their results indicate that insiders' sales which are based on a 10b5-1 plan systematically follow positive and precede negative firm performance, generating abnormal forward-looking returns larger than those earned by insiders' sales that are not based on a plan.

What we want to check first is how the trades that are based on a plan are spread throughout the quarter. For each trade in a quarter the timediff variable shows the number of days past the earnings announcement date, which gives us the opportunity to classify the trades into white, black, and blackest periods. We also want to check whether 10b5-1 participants are more willing than non-participants to trade in a short window immediately before earnings announcement date. The results will clarify both the insiders' behavior and whether they are strategically using 10b5-1 plans to reduce the litigation risk. The next test compares the average cumulative abnormal return generated by the 10b5-1 participant trades and non-participant trades. We test this for white, black, and blackest windows, and by comparing the results we can check whether insiders are strategically timing the 10b5-1 plans to trade more freely in the windows when trading is not ordinarily permitted.

1.5 Results

Table 1.3 presents results for stock price volatility and stock price return for the white and the blackest period. In this table, returns are calculated as $(P_1\text{-}P_0)/P_0$ so the Blackest return is $(P_{59}\text{-}P_{50})/P_{50,}$ and the White return is calculated as $(P_{12}\text{-}P_3)/P_3$. As presented in Table 1.3 average return for blackest period is bigger than the average return for white period. The standard deviation is also larger for the blackest period showing the prices are more volatile in the blackest period. This benchmark result establishes that the value of private information and associated trading profits may be greater in the blackest window. That is, an insider who possesses information about the likely stock price after the earnings information is released will be able to buy (or sell) more profitably in a period where the price is more likely to swing away from this value.

We report the CARs over six different windows of (0; 1), (0; 2), (0; 10), (0, 30), (0, 60), and (0, 90). We also separate the results for purchase and sale. Tables 1.4 and 1.5 reports results for hypothesis 1. Table 1.4 shows the CARs over different event windows for purchase and sale for pre-SOX period. CARs are significantly smaller for shorter periods of time, and as the window gets larger respectively. Results on the purchase side indicate that most of the white period trades generate significantly higher abnormal returns than black period trades. On the sale side, the results are mixed; we see that for the longer widows (e.g. 60 and 90 days) black sales generate larger CAR compared to white sales. The 60-day CAR for the black period is -9.94% which is significantly larger than -8.95% CAR for the White one. This pattern continues to the 90 days window as well (-15.08%

CAR for black; -14.85% for white). These results are consistent with the insider trading literature which argues that market reaction to insider purchases is stronger than the reaction to insider sales. Table 1.5 shows the CARs for different windows within the Post-SOX period. The results are similar to Table 1.4. For the purchase side, white generates significantly larger abnormal returns compared to black. However, for the sell side, all black windows generate larger CARs than white windows. Table 1.4 and 1.5 results cannot confirm the idea that black period trades will generate larger CAR generally, as we find mixed results on sale and purchase sides. Surprisingly, white period purchases create larger CAR even though black periods are characterized by information asymmetry. Why would insiders even trade in the black window if it yields lower CAR and raises suspicion? This made me wonder whether all white periods potentially generate larger CAR or is there a selection bias in choosing which white window to trade in. We design a setting to test hypothesis 2 and determine whether all white periods generate larger CAR, or only the white periods in which insiders trade. We compare the CARs of the white periods that the insiders traded in with the average CARs of non-traded white periods in for each individual insider. Table 1.6 shows the results for hypothesis 2-1. We present the results for pre-SOX and post-SOX periods, and for purchase and sell sides separately. For the pre-SOX period, results indicate that traded white periods generate significantly larger CARs than non-traded white periods for both purchases and sales. For the purchase side, the CAR for the white period that the insider decides to trade in is 9.33% which is significantly larger than 0.22% average CARs of all the other white periods that the insider did not trade in. For the sell side, the CAR of the white period that the insider traded in is -14.86% which is significantly smaller than -

1.46% average CARs of all other white periods that the insider did not trade in. The results hold for the post-SOX period as well. This result indicates that not all white windows are the same, and not all white windows have the potential of generating large CAR. The results show that there is a selection bias in choosing which white window to trade in.

Table 1.7 shows results for hypothesis 3. This test checks whether the CAR would have changed if the insider would have postponed their trades to the white period. For the pre-SOX period, the results are significant for the purchase side, but not for the sell side. Purchase side generates 4.89% CAR, which is significantly larger than 3.20% CAR that they would have earned if they would have waited for the white period. For the post-SOX period, the results are significant for both purchase and sell sides. The results generally indicate that black period is generating larger CAR, and insiders would lose by waiting until the white period to trade.

Table 1.9 shows how 10b5-1 trades are spread through the quarter for each year. On average, 19.51% of 10b5-1 sales are in the white period, and 80.49% are in the black period. If the black period is shortened to the last ten days before the earnings announcement date (blackest period), we observe that on average 34.50% of the 10b5-1 sales are executed on that period. An average of 17.57% of purchases are in the white window, and 82.43% are in the black period. checking the results for the blackest period, we observe that 26.84% of purchases are executed in that period.

We also check whether most trades executed in a ten-day pre-announcement window are based on a 10b5-1 plan. Results show that 62.82% of sale trades executed ten days before the earnings announcement date are made under a 10b5-1 plan.

Results for the average CAR of the trades that are executed under a 10b5-1 plan, and non-plan trades for black, white, and the blackest windows are reported in tables 1.10 and 1.11. Table 1.10 shows that both sales and purchases executed under a 10b5-1 plan generate significantly higher abnormal returns in a 60-day windows for black and blackest periods. 10b5-1 plan trades in black window on average generate 5.87% CAR which is significantly larger than 3.13% which is the average CAR for non-plan trades. For the blackest period, 10b5-1 plan trades in a 60-day window on average generate 5.31% CAR which is significantly larger than 2.56% CAR generated by non-plan trades.

1.6 A closer look at companies with insider trading policies

Companies discuss their insider trading policies and regulations on the code of conduct section of the investor relations, if they have any. Checking this section for different companies we make a list of 250 companies which have some type of insider trading policy disclosed in their code of conduct section.[4] The insider trading data for these companies are downloaded from Thomson Reuters TFN Insiders Data. 134,561 trades are downloaded from Thomson Reuters for 218 companies from 2004 to 2014 with transaction code “S” or “P”. Aligning the trades with the earnings announcement, we identify on which day of the quarter each trade was executed. Trades executed on days 3 to 12 are the white window trades, and the ones executed on days 50 to 59 are the blackest window trades. We check the difference in CAR for sells and purchases of white and blackest windows. The results are presented on table 1.12, and most of the trades, both sells and purchases, for these companies are executed in the white period. There are 36,922 sells executed in the white period, while there are only 4,779 sells in the blackest

[4] List of the companies can be found in Appendix A of chapter 1.

period. We have the same results for purchase side. There are 3,381 trades executed in the white window and 562 in the blackest window. For both sells and purchases, the results indicate that the insiders gain much more abnormal return by trading in the white window compared to the blackest window. These companies are the ones with insider trading policies and the trades executed by the insider are watched more closely by the company in the blackest period. The insiders are not questioned for the white period trades and we again see a selection bias. The insiders pick the white window in which the market fails to fully reflect the private information, and they safely trade in that window and gain more cumulative abnormal return.

1.7 Conclusion

“Blackout” period corporate policies are not enforced by the SEC, but some firms decide to employ them and only let insiders trade in a 10-day window that starts three days after the earnings announcement date. The rationale behind this policy is to prevent insider trading on the days there is more information asymmetry between insiders and other traders. Based on the results I cannot argue that black period trades generate larger CAR on average. On the contrary, white period trades generate higher car than black period trades in certain situations. Still, there is a strong selection bias in choosing which white period to trade in by insiders, and it is clear that not all white periods have the same potential to generate abnormal returns. Results also support the fact that if insiders postpone the execution of their trades to the white window, they lose the information advantage they have over other traders. I also observe that the trades which are executed on the black period or a 10-day period tight before earnings announcement day generate more CAR compared to non-plan trades.

Bibliography

Beaver, W.H. 1968, The Information Content of Annual Earnings Announcements. *Journal of Accounting Research Supplement* 6, 67-92.

Bettis, J.C., Coles, J.L., Lemmon, M.L., 2000. Corporate policies restricting trading by insiders. *Journal of Financial Economics* 57, 191–220.

Brochet, F., 2010. Information Content of Insider Trades before and after the Sarbanes-Oxley Act. *The Accounting Review* 85, 419–446.

Cohen, L., Malloy, C., Pomorski, L., 2012. Decoding Inside Information. *The Journal of Finance* 67, 1009–1043.

Cowan, Arnold R. Eventus 8.0 User's Guide, Standard Edition 2.1. (Cowan Research LC, Ames, Iowa, 2007.)

Fishman, M.J., Hagerty, K.M., 1992. Insider Trading and the Efficiency of Stock Prices. *The RAND Journal of Economics* 23, 106–122.

Jagolinzer, A.D., 2008. SEC Rule 10b5-1 and Insiders' Strategic Trade. *Management Science* 55, 224–239.

Lee, I., Lemmon, M., Li, Y., Sequeira, J.M., 2014. Do voluntary corporate restrictions on insider trading eliminate informed insider trading? *Journal of Corporate Finance* 29, 158–178.

Lintner, J., 1965. The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *The Review of Economics and Statistics* 47, 13–37.

Manne, H.G., 1966. In Defense of Insider Trading. *Harvard Business Review* 44, 113.

Newkirk, T.C., and Robertson, M.A. (1998, September). Insider trading–a US perspective. In Speech at the 16th International Symposium on Economic Crime, Jesus College, Cambridge, England, September (Vol. 19, p. 1998). Retrieved from: https://www.sec.gov/news/speech/speecharchive/1998/spch221.htm

Patell, J., and Wolfson, M., 1981. The Ex Ante and Ex Post Price Effects of Quarterly Earnings Announcements Reflected in Option and Stock Prices. *Journal of Accounting Research* 19,434-58.

________., 1984. The Intraday Speed of Adjustment of Stock Prices to Earnings and Dividend Announcements. *Journal of Financial Economics* 13, 223-52.

Roulstone, D.T., 2003. The Relation between Insider-Trading Restrictions and Executive Compensation. *Journal of Accounting Research* 41, 525–551.

Sen, R., 2008. Are Insider Sales Under 10b5-1 Plans Strategically Timed? (SSRN Scholarly Paper No. ID 1154837). Social Science Research Network, Rochester, NY.

Seyhun, H.N., 1992. Effectiveness of the Insider-Trading Sanctions, The. J.L. & Econ. 35, 149.

Sharpe, W.F., 1964. Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk. *The Journal of Finance* 19, 425–442.

Sharpe, W.F., 1963. A Simplified Model for Portfolio Analysis. *Management Science* 9, 277–293.

Shon, J., Veliotis, S., 2013. Insiders' Sales Under Rule 10b5-1 Plans and Meeting or Beating Earnings Expectations. *Management Science* 59, 1988–2002.

1.8 Tables for Chapter 1

Table 1.1 Summary statistics

	Total Trades	Form 4s	OMP	OMS	# Shares Traded	$ Value Traded	Shares/ Outstanding	Mkt cap%
1998	552,179	256,870	63,681	73,139	2,940,939,186	94,393,996,321	1.396	1.019
1999	525,866	254,124	63,877	75,725	3,131,548,962	96,749,582,495	1.275	0.852
2000	556,960	285,372	58,227	109,101	3,438,759,009	121,968,837,773	1.133	0.806
2001	488,476	260,712	44,857	115,946	3,680,073,590	69,774,758,313	1.009	0.516
2002	455,157	280,871	59,685	114,081	2,595,627,054	38,973,047,208	0.686	0.323
2003	582,781	349,014	35,191	182,605	4,524,377,567	60,664,375,155	1.226	0.635
2004	636,868	407,531	34,927	225,995	5,589,925,515	73,640,889,184	1.444	0.590
2005	685,309	473,355	43,642	275,979	4,637,002,361	77,557,948,997	1.140	0.553
2006	738,023	534,289	43,775	335,418	3,880,285,940	83,287,778,224	0.921	0.567

This table presents summary statistics of the insider trading activity over the period of 1998 to 2006. Insider trade data is from the Thomson Reuters database. “Total Trades” shows the total number of trades executed by insiders. “Cleansed” shows the number of trades that are flagged as cleansed (reliable) by the Thomson Reuters database. The next column shows the number of cleansed form 4s filed by insiders. OMP and OMS are the open market purchases and sales executed by insiders. “# Shares Traded” shows the total number of shares trades by insiders in open market sells and purchases. “$ Value Traded” shows the total dollar amount of all the open market sales and purchases executed by insiders. “Mkt cap%” is total dollar value of all open market sells and purchases executed by insiders divided by market capitalization (total market value of the shares outstanding.)

Table 1.2 Summary Statistics of Insider Trades by Their Roles

	CEO	CEO B	CEO W	Total/ B	CFO	CFO B	CFO W	Total/ B	CB	CB B	CB W	Total/ B
1998	12723	7839	4884	1.61	4504	2733	1771	1.54	3655	2342	1313	1.78
	12.19%	11.36%	13.82%		4.32%	3.96%	5.01%		3.50%	3.39%	3.72%	
1999	14382	8769	5613	1.56	4913	3118	1795	1.74	3063	2007	1056	1.9
	13.31%	16.16%	11.96%		4.55%	4.25%	5.17%		2.83%	2.74%	3.04%	
2000	15927	10328	5599	1.84	6597	4057	2540	1.6	4332	2931	1401	2.09
	12.21%	11.86%	12.92%		5.06%	4.66%	5.86%		3.32%	3.37%	3.23%	
2001	17801	12118	5683	2.13	5520	3500	2020	1.73	5957	3686	2277	1.62
	13.77%	9.37%	13.63%		4.27%	3.99%	4.84%		4.61%	4.20%	5.46%	
2002	21205	13966	7239	1.93	5265	3310	1955	1.69	7048	5186	1862	2.79
	14.42%	9.54%	4.95%		3.60%	2.26%	1.34%		4.82%	5.13%	4.11%	
2003	28133	20307	7826	2.59	9334	6308	3026	2.08	7070	4682	2388	1.96
	15.21%	15.98%	13.69%		5.05%	4.94%	5.29%		3.82%	3.66%	4.18%	
2004	38252	26268	11984	2.19	10752	6871	3881	1.77	7003	4846	2157	2.25
	17.18%	17.52%	16.49%		4.83%	4.58%	5.34%		3.15%	3.23%	2.97%	
2005	68153	48026	20127	2.39	14234	8717	5517	1.58	7959	5158	2801	1.84
	24.52%	25.70%	22.10%		5.12%	4.66%	6.06%		2.86%	2.76%	3.08%	
2006	89652	64289	25363	2.53	18049	12221	5828	2.1	11325	7217	4108	1.76
	27.20%	28.70%	24.02%		5.48%	5.46%	5.52%		3.44%	3.22%	3.89%	

Table 1.3 Comparison Between Stock Price Return and Volatility for the White and the Blackest Windows

Variable	**Mean**	**Std. Err.**	**Std. Dev**	**[95% Conf. Interval]**	
BlacstRet	0.13266	0.00574	0.4258803	0.1214092	0.1439164
WhiteRet	0.08423	0.00157	0.1707162	0.081141	0.0873334
Combined	0.09974	0.00213	0.2800053	0.0955602	0.1039336
diff	0.04842	0.0059		0.0367541	0.0600971

diff= mean (BlackestRet) - mean (WhiteRet) t=8.1335
Ho : diff = 0

Ha:diff<0	Ha: diff != 0	Ha: diff > 0
Pr (T < t) =1.0000	Pr (|T|< |t|) =0.0000	Pr (T > t) =0.0000

Table 1.4 Average Cumulative Abnormal Return of Sales and Purchases for pre-SOX Period

Panel A

Purchase						
	(0; 1)	(0; 2)	(0; 10)	(0; 30)	(0; 60)	(0; 90)
Black	0.79%	1.01%	2.31%	5.04%	7.13%	8.11%
White	1.42%	1.59%	2.71%	4.61%	8.57%	9.70%
Difference	0.63%	0.58%	0.40%	0.43%	-1.44%	-1.59%
t-value	12.64	10.65	4.64	-3.03	7.18	8.11

Panel B

Sale						
	(0; 1)	(0; 2)	(0; 10)	(0; 30)	(0; 60)	(0; 90)
Black	0.40%	0.24%	-0.92%	-4.49%	-9.94%	-15.08%
White	0.21%	0.01%	-1.25%	-4.68%	-8.95%	-14.85%
Difference	0.19%	0.23%	0.33%	0.19%	-0.99%	-0.23%
t-value	-5.69	-6.52	-6.03	-1.98	7.36	2.34

This table reports the cumulative abnormal return of insider trades for the years 1998 to 2002. Panel A shows the result for purchases and panel B shows the results for sells.

Table 1.5 Average Cumulative Abnormal Return of Sales and Purchases for post-SOX Period

Panel A

Purchase						
	(0; 1)	(0; 2)	(0; 10)	(0; 30)	(0; 60)	(0; 90)
Black	1.07%	1.62%	3.02%	2.93%	3.10%	2.55%
White	1.55%	2.38%	4.14%	5.50%	6.23%	7.13%
Difference	-0.48%	-0.76%	-1.12%	-2.57%	-3.13%	-4.58%
t-value	9.68	12.79	10.47	14.47	10.54	11.89

Panel B

Sale						
	(0; 01)	(0; 2)	(0; 10)	(0; 30)	(0; 60)	(0; 90)
Black	-0.24%	-0.33%	-0.97%	-3.02%	-6.38%	-9.32%
White	-0.21%	-0.32%	-1.24%	-2.71%	-5.14%	-8.11%
Difference	-0.03%	-0.01%	0.27%	-0.31%	-1.24%	-1.21%
t-value	-23.86	-19.37	-14.15	9.56	24.42	17.34

This table reports the cumulative abnormal return of insider trades for the years 2003 to 2006. Panel A shows the result for purchases and panel B shows the results for sales.

Table 1.6 Comparison between Traded and non-traded White Windows

Panel A

Pre SOX Period	**Purchase**	**Sale**
Number of Trades	34,595	122,714
Average CAR of the Traded White Period	9.33%	-14.86%
Average CAR of other White Periods	0.22%	-1.46%
difference	9.11%	-13.40%
t-value	41.22	-99.04

Panel B

Post SOX Period	Purchase	Sale
Number of Trades	50,860	267,800
Average CAR of the Traded White Period	7.15%	-6.08%
Average CAR of other White Periods	0.51%	-1.26%
difference	6.64%	-4.82%
t-value	30.71	-24.45

This table reports a comparison between the average cumulative abnormal return of the white period in which the insider trades and the average cumulative abnormal return of all the other white periods that the insider does not trade in.

Table 1.7 Average CAR of Black window trades postponed to white period

Panel A

Pre Sox Period	**Purchase**	**Sale**
Number of Trades	17,644	25,781
Black Period Trade CAR	4.89%	-13.11%
Forwarded to White CAR	3.20%	-12.48%
difference	1.69%	-0.63%
t-value	3.41	1.53

Panel B

Post Sox Period	**Purchase**	**Sale**
Number of Trades	14,179	77,164
Black Period Trade CAR	4.89%	-10.50%
Forwarded to White CAR	3.20%	-9.61%
difference	1.69%	-0.89%
t-value	2.98	5.41

This table reports the average cumulative abnormal return for the trades that are forwarded from the black period to the white period. Panel A shows the results for the period of 1998 to 2002, and Panel B shows the results for the period of 2003 to 2006.

Table 1.8 Summary Statistics of the 10b5-1 Trades

Panel A

	Total Trades	**OMP**	**OMS**	**10b5-1 S**
2004	85,528	21,862	43,781	19,763
2005	81,796	22,300	40,311	19,044
2006	81,182	20,085	36,839	24,119
2007	87,669	26,068	33,998	26,747
2008	72,413	31,366	21,765	16,203
2009	52,850	21,627	18,074	12,555
2010	61,449	19,101	22,007	20,095
2011	52,645	14,976	19,203	18,107
2012	49,614	11,908	19,488	17,770
2013	56,937	9,780	22,902	23,659
2014	54,833	11,211	18,988	23,770

Panel B

	Merged with CRSP&COMPUSTAT			
	OMP	**OMS**	**10b5-1 S**	**10b5-1 P**
2004	11,640	34,711	14,312	104
2005	13,313	30,629	16,521	122
2006	11,652	29,182	18,837	112
2007	15,858	25,297	23,073	766
2008	22,245	17,216	14,223	2,474
2009	14,981	14,281	10,981	416
2010	11,244	18,090	16,859	189
2011	10,582	15,467	14,005	242
2012	8,237	17,055	11,918	340
2013	6,216	19,675	14,757	461
2014	3,099	8,779	7,919	277

Table 1.9 Summary Statistics of the 10b5-1 Trades

Year	10b5-1 S					
	White		**Black**		**Blackest**	
2004	2,700	18.87%	11,612	81.13%	4,766	33.30%
2005	3,138	18.99%	13,383	81.01%	5,398	32.67%
2006	3,886	20.63%	14,951	79.37%	6,268	33.27%
2007	4,492	19.47%	18,581	80.53%	8,061	34.94%
2008	2,821	19.83%	11,402	80.17%	4,636	32.60%
2009	2,043	18.60%	8,938	81.40%	3,667	33.39%
2010	3,209	19.03%	13,650	80.97%	5,967	35.39%
2011	2,908	20.76%	11,097	79.24%	5,276	37.67%
2012	2,277	19.11%	9,641	80.89%	4,012	33.66%
2013	3,074	20.83%	11,683	79.17%	5,170	35.03%
2014	1,467	18.53%	6,452	81.47%	2,972	37.53%
		19.51%		80.49%		34.50%

Year	10b5-1 P					
	White		**Black**		**Blackest**	
2004	12	11.54%	92	88.46%	33	31.73%
2005	25	20.49%	97	79.51%	31	25.41%
2006	11	9.82%	101	90.18%	35	31.25%
2007	50	6.53%	716	93.47%	199	25.98%
2008	355	14.35%	2,119	85.65%	606	24.49%
2009	83	19.95%	333	80.05%	152	36.54%
2010	51	26.98%	138	73.02%	47	24.87%
2011	66	27.27%	176	72.73%	58	23.97%
2012	76	22.35%	264	77.65%	77	22.65%
2013	72	15.62%	389	84.38%	95	20.61%
2014	51	18.41%	226	81.59%	77	27.80%
		17.57%		82.43%		26.84%

Table 1.10 Average Cumulative Abnormal Return of 10b5-1 Plans for pre-SOX Period

Panel A

White Period Sell				
	(0; 10)	(0; 30)	(0; 60)	(0; 90)
0	-0.84%	-1.80%	-4.00%	-5.81%
1	-0.68%	-1.72%	-4.18%	-5.89%
Difference	-0.16%	-0.08%	0.18%	0.08%
t-value	-3.57	-1.07	1.43	-0.45

Panel B

Black Period Sell				
	(0; 10)	(0; 30)	(0; 60)	(0; 90)
0	-0.67%	-2.04%	-4.05%	-5.92%
1	-0.46%	-1.99%	-4.06%	-6.36%
Difference	-0.21%	-0.05	0.01	0.44%
t-value	-7.72	-1.06	0.07	4.23

Panel C

Blackest Period Sell				
	(0; 10)	(0; 30)	(0; 60)	(0; 90)
0	-0.70%	-2.05%	-3.63%	-5.72%
1	-0.28%	-1.69%	-3.60%	-6.21%
Difference	-0.42%	-0.36%	-0.03%	0.49%
t-value	-6.36	-3.17	-0.19	2.22

Panels A, B, and C report results for white, black, and blackest windows respectively. We do not see significant results for white period, but black and blackest period results indicate that 10b5-1 sales generate significantly larger CAR.

Table 1.11 Average Cumulative Abnormal Return of 10b5-1 Plans for post-SOX Period

Panel A

White Period Purchase				
	(0; 10)	(0; 30)	(0; 60)	(0; 90)
0	3.05%	4.12%	5.30%	6.22%
1	2.71%	3.54%	1.81%	1.06%
Difference	0.34%	0.58%	3.49%	5.16%
t-value	0.76	1.21	7.18	7.45

Panel B

Black Period Purchase				
	(0; 10)	(0; 30)	(0; 60)	(0; 90)
0	2.08%	2.73%	3.13%	3.65%
1	1.66%	3.82%	5.87%	8.30%
Difference	0.42%	-1.09%	-2.74%	-4.65%
t-value	2.91	-3.66	-7.39	-8.96

Panel C

Blackest Period Purchase				
	(0; 10)	(0; 30)	(0; 60)	(0; 90)
0	1.52%	1.60%	2.56%	2.36%
1	0.07%	1.29%	5.31%	3.29%
Difference	1.45%	0.31%	-2.75%	-0.93%
t-value	6.89	0.93	-5.48	-1.32

Table 1.12 Average Cumulative Abnormal Return of Sales and Purchases for Companies with Insider Trading Policy

Panel A

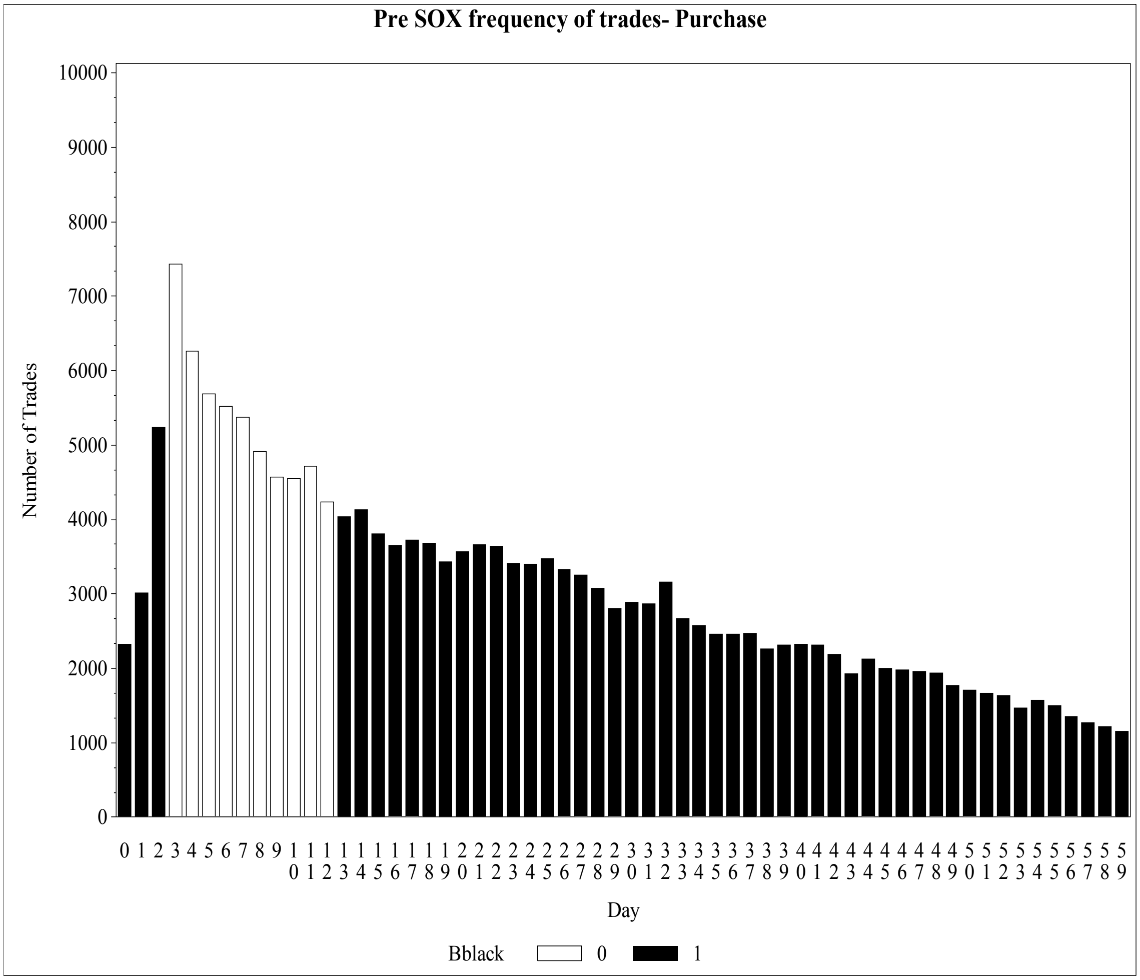
Sale							
	Number of Trades	(0; 1)	(0; 2)	(0; 10)	(0; 30)	(0; 60)	(0; 90)
Blackest	4,779	-0.20%	0.00%	-0.19%	-0.56%	-1.56%	-3.06%
White	36,922	-0.27%	-0.29%	-1.60%	-4.01%	-5.05%	-6.54%

Panel B

Purchas							
	Number of Trades	(0; 1)	(0; 2)	(0; 10)	(0; 30)	(0; 60)	(0; 90)
Blackest	562	0.60%	1.37%	2.01%	0.87%	3.52%	2.14%
White	3,381	1.24%	2.13%	4.02%	6.66%	10.50%	12.71%

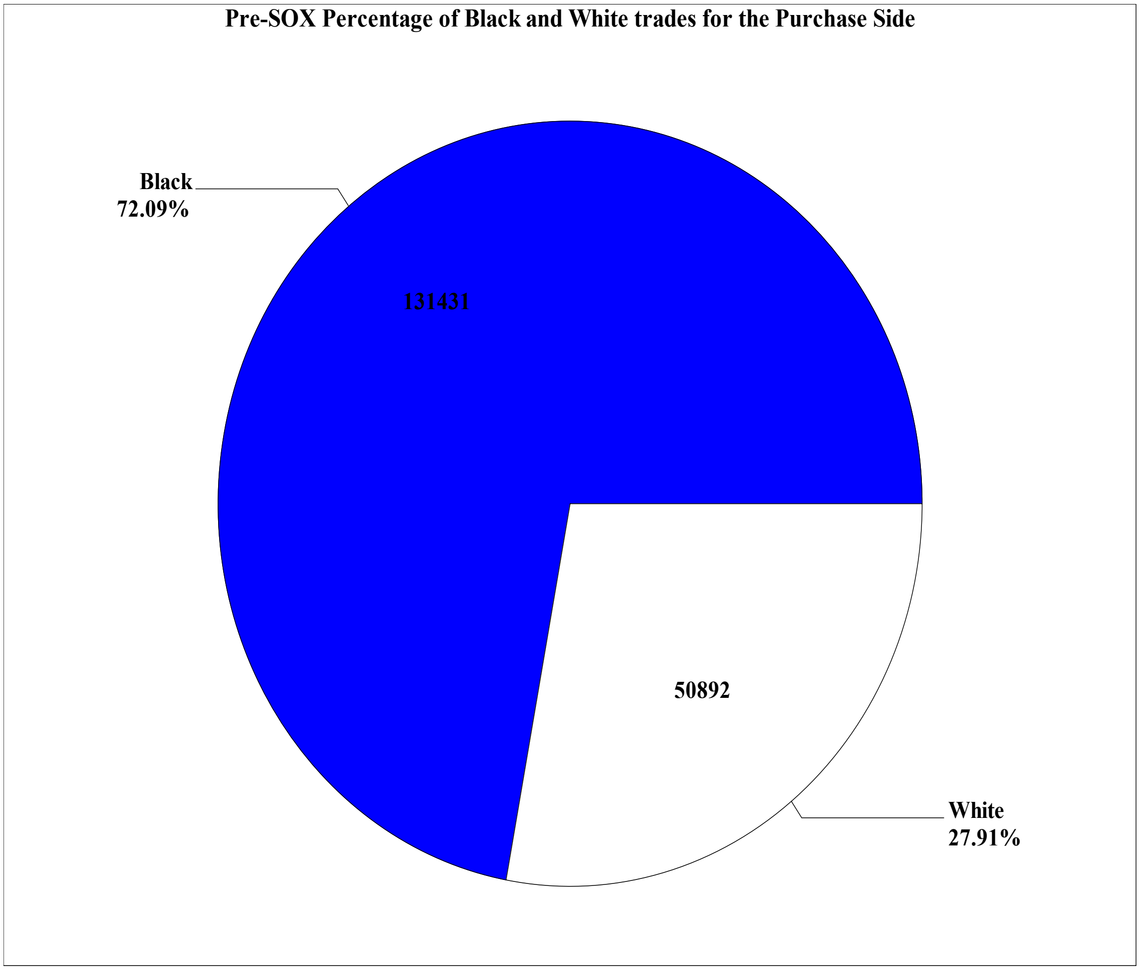
1.9 Figures for Chapter 1

Figure 1.1 Frequency of Purchases for the pre-SOX Period



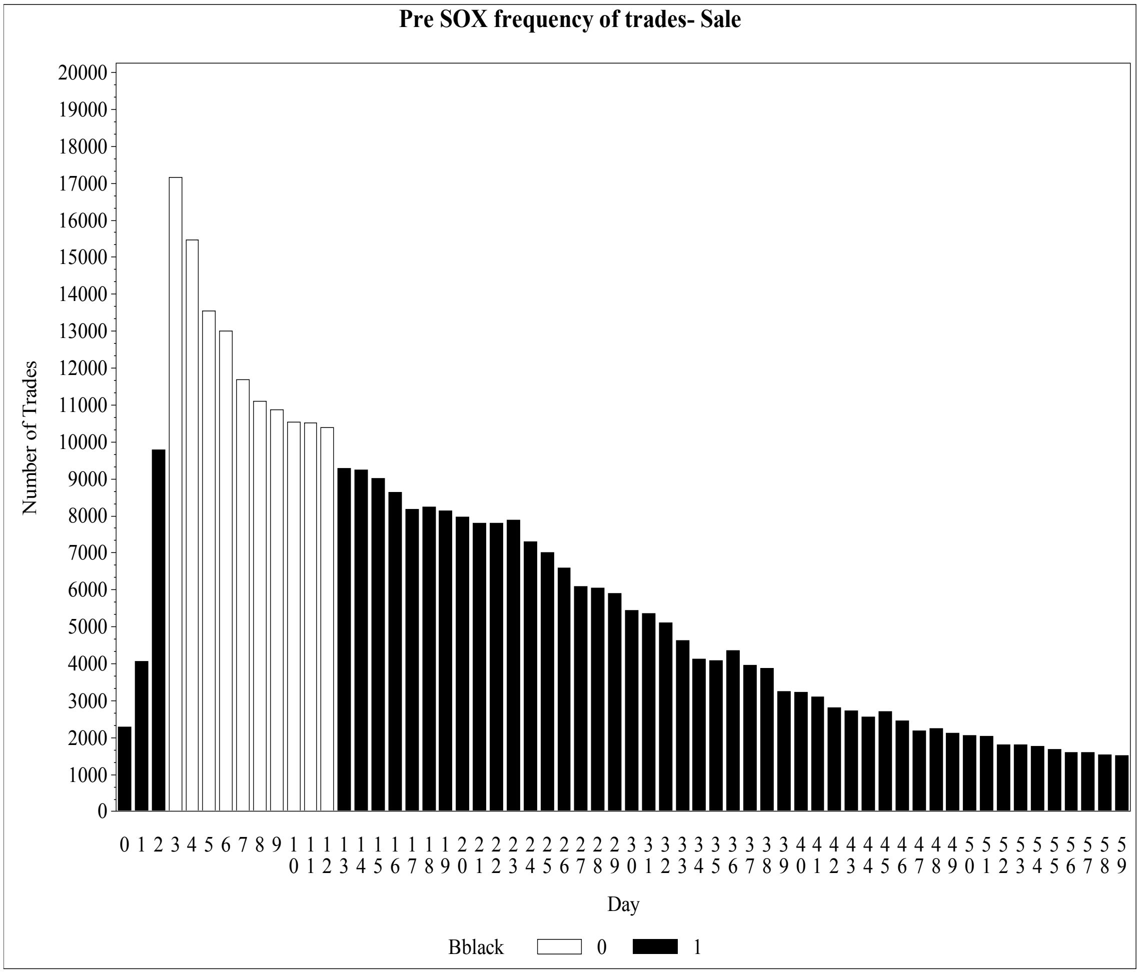
This figure shows how trades are spread throughout the quarter. It specifically shows the frequency of purchases for the pre-SOX period. White trades are executed three to 12 days after the earnings announcement date which are shown in white. The rest of the trades are the ones executed in the black period, which are shown in black.

Figure 1.2 Percentage of Black and White Purchases for the pre-SOX Period



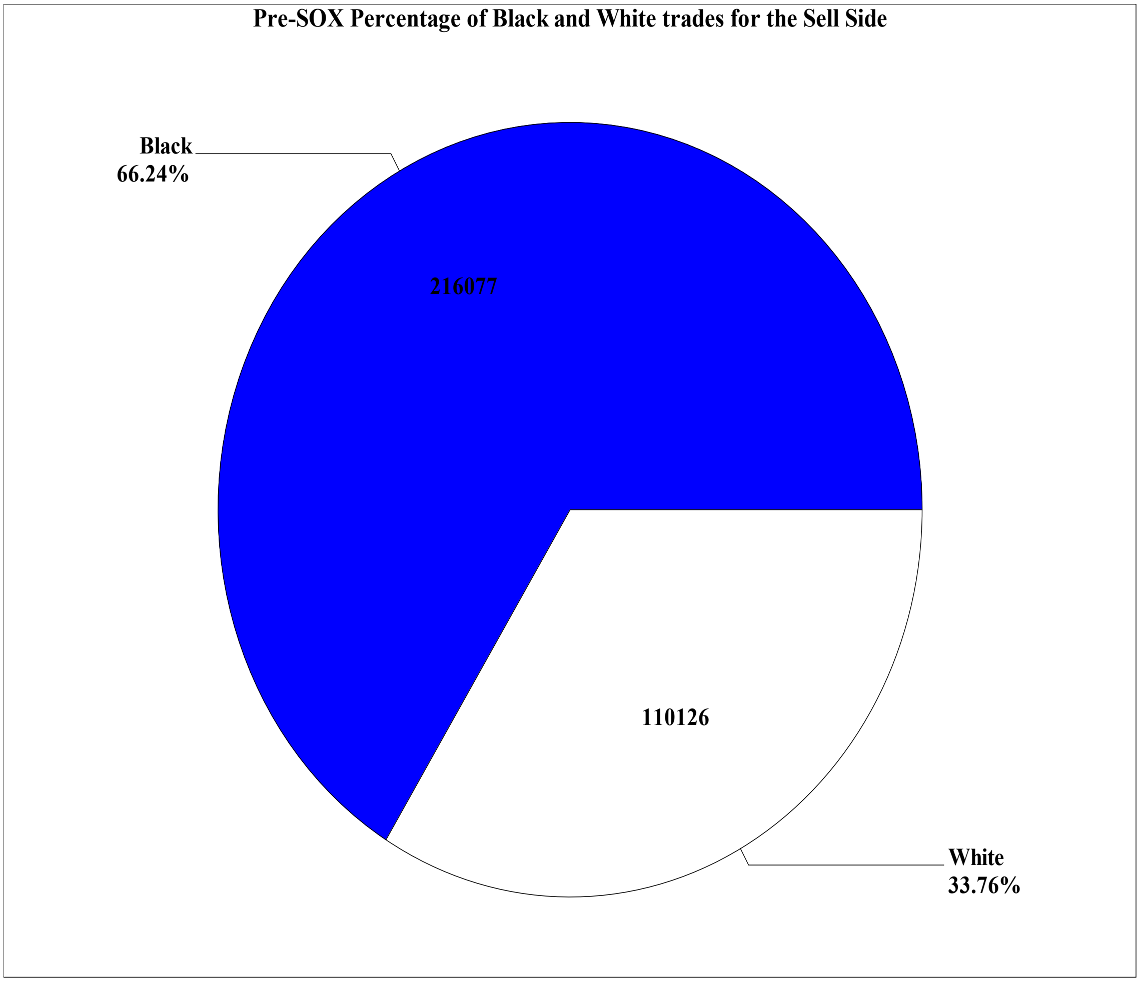
This figure shows the the total number and the percentage of purchases that are executed in the white and black period. For the pre-SOX period, 72.09% of the purchases are executed in the black period and 27.91% of the purchases are executed in the white period.

Figure 1.3 Frequency of Sells for the pre-SOX Period



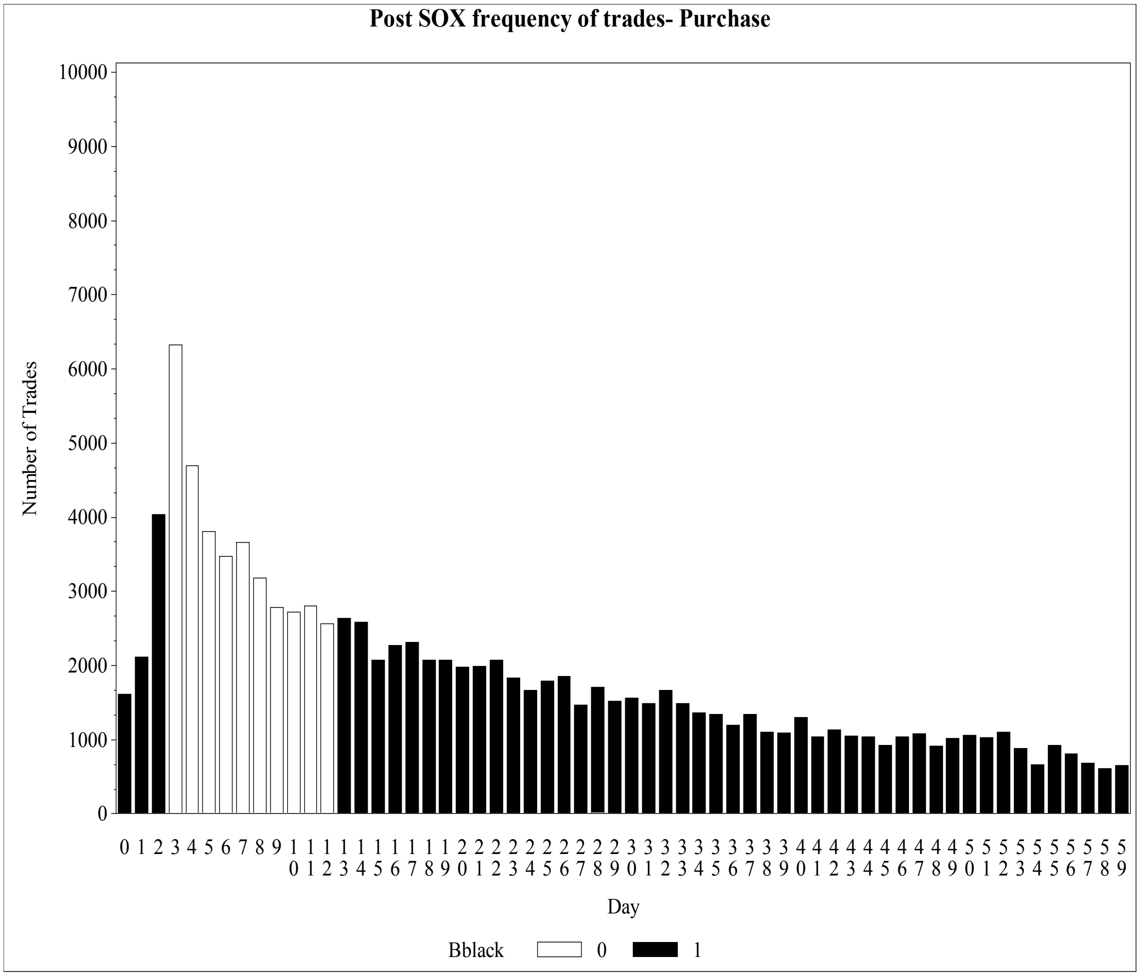
This figure shows how trades are spread throughout the quarter. It specifically shows the frequency of sales for the pre-SOX period. White trades are the ones that are executed three to 12 days after the earnings announcement date which are shown in white. The rest of the trades are the ones executed in the black period, which are shown in black.

Figure 1.4 Percentage of Black and White Sells for the pre-SOX Period



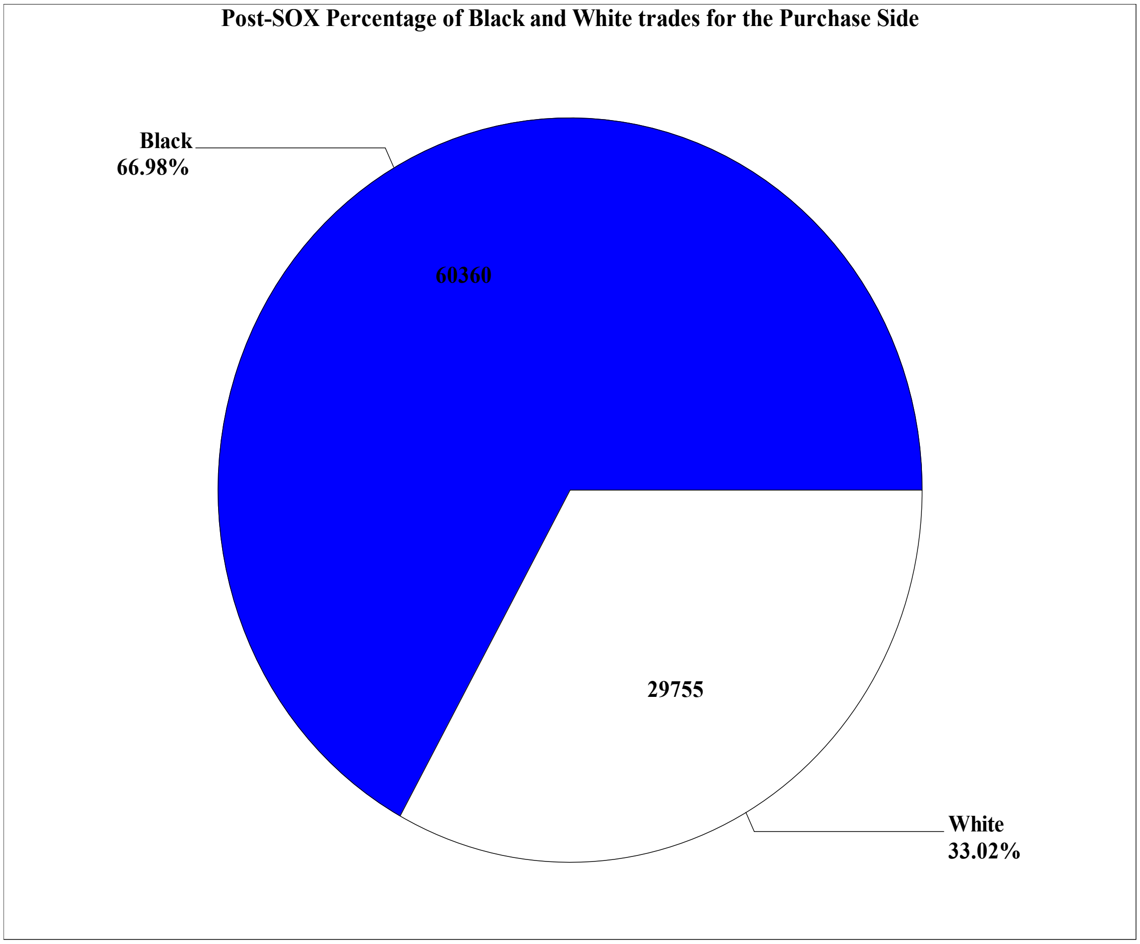
This figure shows the the total number and the percentage of sales executed in the white and black period. For the pre-SOX period 66.24% of the sales are executed in the black period and 33.76% of the sales are executed in the white period.

Figure 1.5 Frequency of Purchases for the post-SOX Period



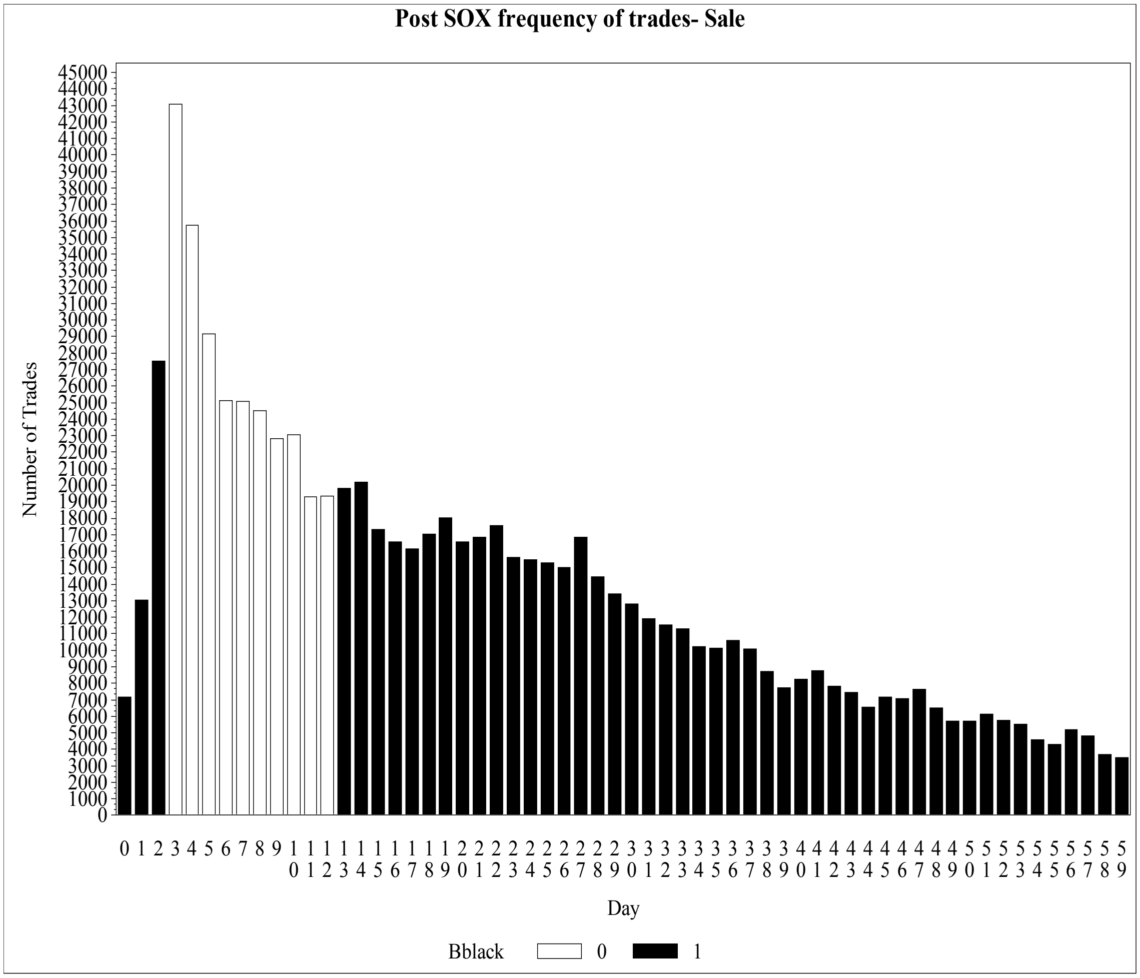
This figure shows how trades are spread throughout the quarter. It specifically shows the frequency of purchases for the post-SOX period. White trades are the ones that are executed three to 12 days after the earnings announcement date which are shown in white. The rest of the trades are the ones executed in the black period, which are shown in black.

Figure 1.6 Percentage of Black and White Purchases for the post-SOX Period



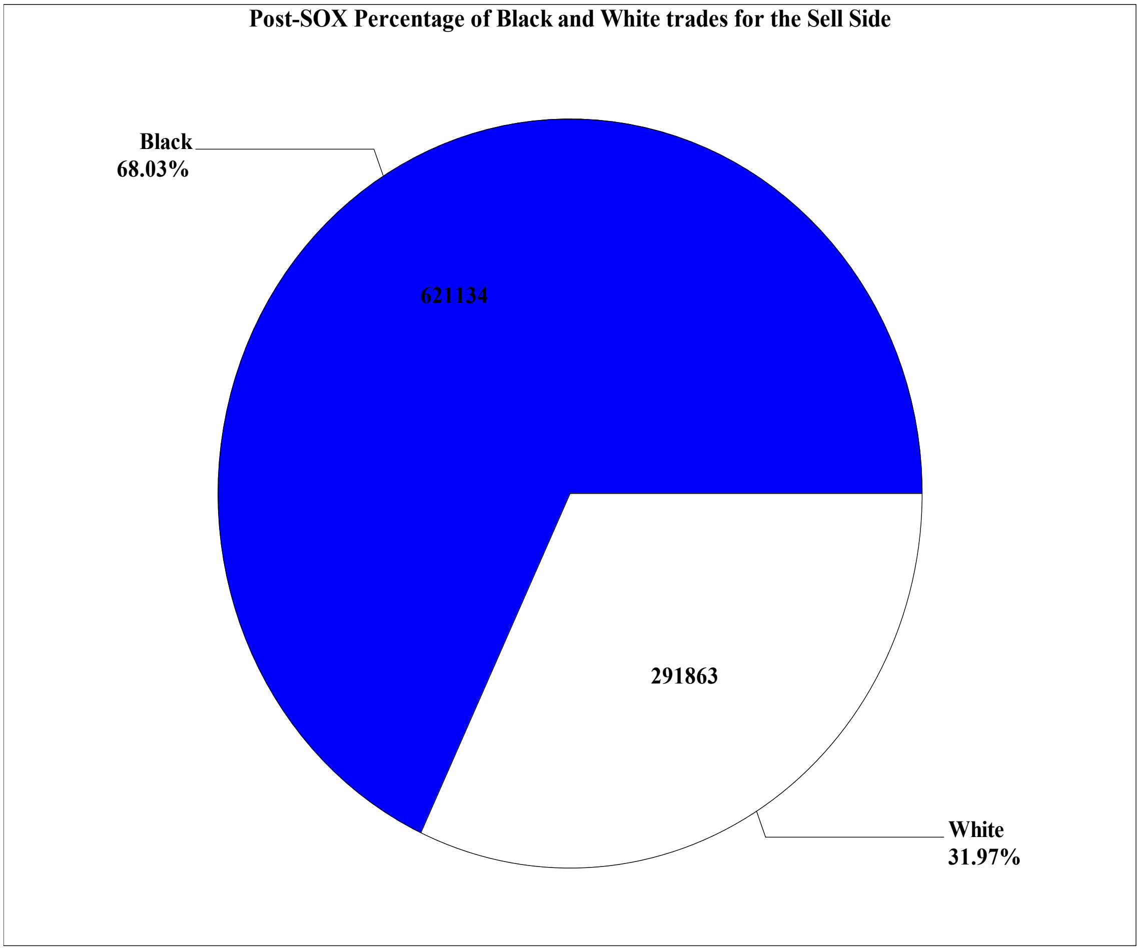
This figure shows the the total number and the percentage of purchases executed in the white and black period. For the post-SOX period 66.98% of the purchases are executed in the black period and 33.02% of the purchases are executed in the white period.

Figure 1.7 Frequency of Sells for the post-SOX Period



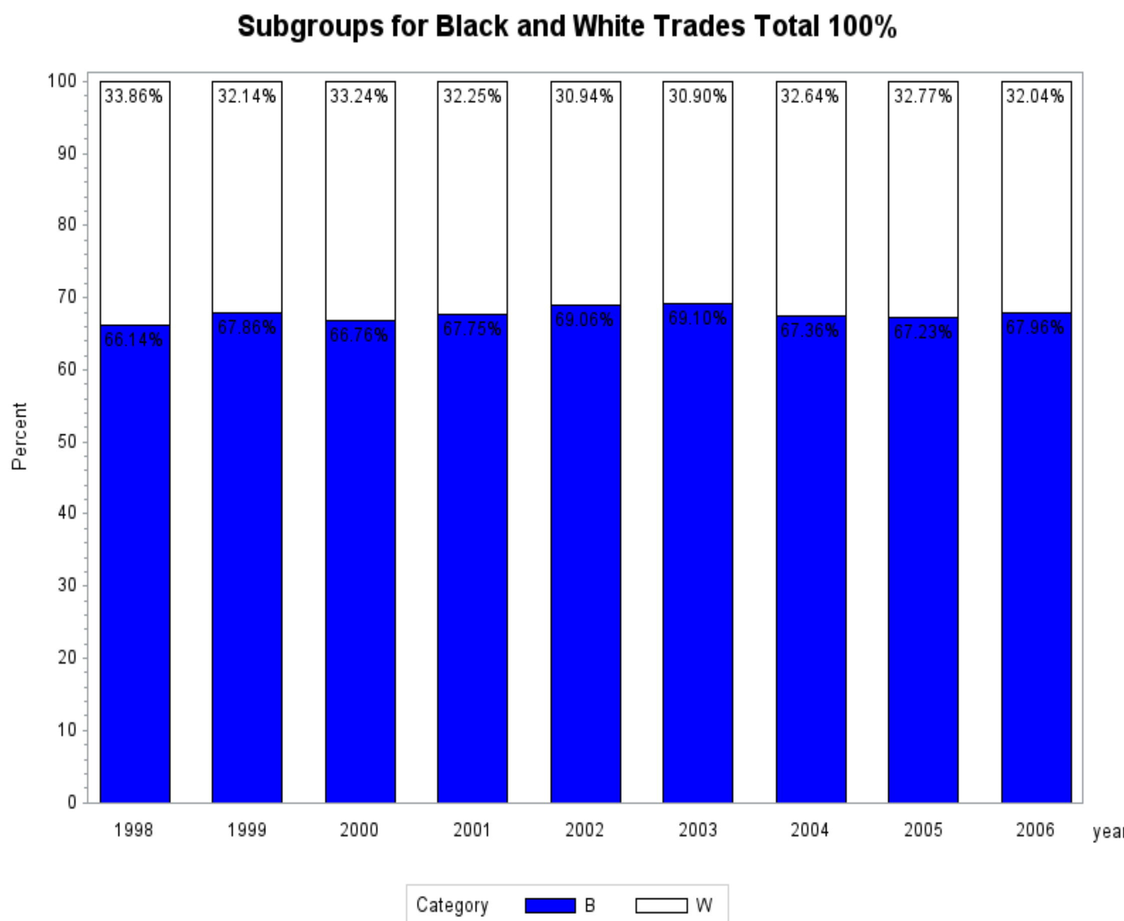
This figure shows how trades are spread throughout the quarter. It specifically shows the frequency of sales for the post-SOX period. White trades are the ones that are executed three to 12 days after the earnings announcement date which are shown in white. The rest of the trades are the ones executed in the black period, which are shown in black.

Figure 1.8 Percentage of Black and White Sells for the post-SOX Period



This figure shows the the total number and the percentage of sales that are executed in the white and black period. For the post-SOX period 68.03% of the sales are executed in the black period and 31.97% of the sales are executed in the white period.

Figure 1.9 Annual Percentage of Black and White Trades



This figure shows the percentage of trades executed in the black and white windows throughout our sample period. The trades that are executed in the black period are shown in blue and the trades that are executed in the white period are shown in white. On average 67.69% of all open market sales and purchases executed by the insiders are in the black window.

1.10 Appendix for Chapter 1

1.10.1 Appendix A List of Companies with Insider Trading Policy

AAR CORP	CTS CORP	ENZON PHARMACEUTICALS INC
PINNACLE WEST CAPITAL CORP	CALAMP CORP	ESCALADE INC
ABBOTT LABORATORIES	CAMPBELL SOUP CO	FAMILY DOLLAR STORES
ALASKA AIR GROUP INC	CONSTELLATION BRANDS	FIRSTMERIT CORP
ARCONIC INC	CARDINAL HEALTH INC	TRUSTMARK CORP
ABM INDUSTRIES INC	CARPENTER TECHNOLOGY CORP	M & T BANK CORP
AMERICAN EXPRESS CO	CANTEL MEDICAL CORP	FLOW INTL CORP
AFLAC INC	CHEMED CORP	FRANKLIN ELECTRIC CO INC
AMERICAN GREETINGS -CL A	CHEVRON CORP	GAP INC
AMERICAN NATIONAL INSURANCE	CHUBB CORP	GENERAL ELECTRIC CO
AMERICAN VANGUARD CORP	CHURCH & DWIGHT INC	GENERAL MILLS INC
ANAREN INC	CINCINNATI BELL INC	SPX CORP
APOGEE ENTERPRISES INC	CLOROX CO/DE	GENUINE PARTS CO
APPLE INC	COCA-COLA CO	GRACO INC
SOUTHERN GAS CO	COLGATE-PALMOLIVE CO	GRAHAM CORP
ATWOOD OCEANICS	COMERICA INC	GRAINGER (W W) INC
AUTOMATIC DATA PROCESSING	COMPUTER SCIENCES CORP	HEICO CORP
AVERY DENNISON CORP	CONAGRA FOODS INC	HOME DEPOT INC
AVNET INC	CMS ENERGY CORP	HNI CORP
AVON PRODUCTS	COOPER TIRE & RUBBER CO	HOOPER HOLMES INC
BALL CORP	COURIER CORP	HORMEL FOODS CORP
BAXTER INTERNATIONAL INC	CRANE CO	CENTERPOINT ENERGY INC
BECTON DICKINSON & CO	CRAWFORD & CO	URSTADT BIDDLE PROPERTIES
VERIZON COMMUNICATIONS INC	CROWN HOLDINGS INC	HURCO COMPANIES INC
BOLT TECHNOLOGY CORP	CUMMINS INC	IMMUNOMEDICS INC

BRADY CORP	DEERE & CO	GRIFFON CORP
BRISTOL-MYERS SQUIBB CO	DELUXE CORP	INTEGRATED DEVICE TECH INC
CALERES INC	DTE ENERGY CO	INTEL CORP
MATERION CORP	DILLARDS INC -CL A	INTERFACE INC
MASTEC INC	DIODES INC	INTL FLAVORS & FRAGRANCES
CIGNA CORP	DISNEY (WALT) CO	NAVISTAR INTERNATIONAL CORP
CSS INDUSTRIES INC	DOVER CORP	INTL PAPER CO
CNA FINANCIAL CORP	OMNICOM GROUP	INVACARE CORP
CSX CORP	PERKINELMER INC	KAYDON CORP

KAYDON CORP	TENET HEALTHCARE CORP	QUESTAR CORP
KIMBERLY-CLARK CORP	NATIONAL PRESTO INDS INC	RPM INTERNATIONAL INC
LSB INDUSTRIES INC	NEW YORK TIMES CO -CL A	RITE AID CORP
LSI CORP	NEWELL BRANDS INC	ROYAL GOLD INC
LEE ENTERPRISES INC	NEWMONT MINING CORP	SEI INVESTMENTS CO
LEGGETT & PLATT INC	NIKE INC	ST JUDE MEDICAL INC
LEGG MASON INC	NORDSON CORP	SCANA CORP
L BRANDS INC	NORFOLK SOUTHERN CORP	SHERWIN-WILLIAMS CO
LINCOLN NATIONAL CORP	NACCO INDUSTRIES -CL A	SYMMETRICOM INC
LOCKHEED MARTIN CORP	NORTHERN TRUST CORP	SMITHFIELD FOODS INC
LOEWS CORP	TEREX CORP	SONOCO PRODUCTS CO
LOUISIANA-PACIFIC CORP	WELLS FARGO & CO	PIONEER ENERGY SERVICES CORP
LOWE'S COMPANIES INC	FIRSTENERGY CORP	EDISON INTERNATIONAL
LUBYS INC	ONE LIBERTY PROPERTIES INC	AMERICAN STATES WATER CO
MDU RESOURCES GROUP INC	OWENS & MINOR INC	SOUTHERN CO
MTS SYSTEMS CORP	PNC FINANCIAL SVCS GROUP INC	SOUTHWEST AIRLINES
MASCO CORP	PACCAR INC	SPARTAN MOTORS INC
MATERIAL SCIENCES CORP	PARKER-HANNIFIN CORP	STANDARD MOTOR PRODS

MAUI LAND & PINEAPPLE CO	PARKWAY PROPERTIES INC	PUBLIC STORAGE
MAXWELL TECHNOLOGIES INC	AMERICAN FINANCIAL GROUP INC	SUPERIOR INDUSTRIES INTL
MCDERMOTT INTL INC	PPL CORP	SWIFT ENERGY CO
MCDONALD'S CORP	PENNSYLVANIA RE INVS TRUST	RS LEGACY CORP
MEDICAL ACTION INDUSTRIES	PEPSICO INC	ALLEGHENY TECHNOLOGIES INC
CVS HEALTH CORP	AQUA AMERICA INC	TESORO CORP
MENTOR GRAPHICS CORP	PVH CORP	TEXAS INSTRUMENTS INC
MEREDITH CORP	PIEDMONT NATURAL GAS CO	THERMO FISHER SCIENTIFIC INC
MICROS SYSTEMS INC	AGILYSYS INC	TIMKEN CO
ENTERGY CORP	PITNEY BOWES INC	TOOTSIE ROLL INDUSTRIES INC
MSA SAFETY INC	PEPCO HOLDINGS INC	TORCHMARK CORP
3M CO	PROTECTIVE LIFE CORP	TORO CO
RUBY TUESDAY INC	PUBLIC SERVICE ENTRP GRP INC	TRINITY INDUSTRIES
MURPHY OIL CORP	PULTEGROUP INC	UNION PACIFIC CORP
MYERS INDUSTRIES INC	QUAKER CHEMICAL CORP	UDR INC
NATIONAL FUEL GAS CO	QUALITY SYSTEMS INC	SPRINT CORP

UNIVERSAL CORP/VA	FAIR ISAAC CORP
VALSPAR CORP	AKORN INC
VISHAY INTERTECHNOLOGY INC	WASTE MANAGEMENT INC
WGL HOLDINGS INC	BMC SOFTWARE INC
AVISTA CORP	1ST SOURCE CORP
WEINGARTEN REALTY INVST	TCF FINANCIAL CORP
WEIS MARKETS INC	ARGO GROUP INTL HOLDINGS LTD
WEYERHAEUSER CO	MERCURY GENERAL CORP
ALLIANT ENERGY CORP	WESBANCO INC
FOOT LOCKER INC	ROYAL BANCSHARES/PA -CL A
XEROX CORP	FIRST FINL BANCORP INC/OH

WHITE MTNS INS GROUP LTD	TOMPKINS FINANCIAL CORP
ASSOCIATED BANC-CORP	FARMERS CAPITAL BANK CORP
FIRST MIDWEST BANCORP INC	CAPITAL SOUTHWEST CORP
AEP INDUSTRIES INC	YUMA ENERGY INC -OLD
EMC CORP/MA	GANNETT CO INC
BIG LOTS INC	INVESTMENT TECHNOLOGY GP INC
ORACLE CORP	HOST HOTELS & RESORTS LP
SIGMA DESIGNS INC	RLI CORP
AMAG PHARMACEUTICALS INC	GENERAL COMMUNICATION -CL A
WERNER ENTERPRISES INC	CALGON CARBON CORP
HARLEY-DAVIDSON INC	BERRY PETROLEUM -CL A
TRANS WORLD ENTMT CORP	CONMED CORP
WATTS WATER TECHNOLOGIES INC	SYNOVUS FINANCIAL CORP
VALHI INC	PROGRESSIVE CORP-OHIO
RENTRAK CORP	ORASURE TECHNOLOGIES INC

Chapter 2: The Ability of the Banking Sector to Predict Financial Crisis Using Text Analytics

2.1 Introduction

The U.S. Securities and Exchange Commission (SEC) requires all of its registrants to submit annual and quarterly reports (which include the registrant's financial statements) as well as other documents such as press releases. The annual report issued by the firm and published by the SEC is considered an important and detailed source of information for stockholders and investors to procure information about the firms' business, financial statement, and facing risks. Prior research has evaluated to what extent the information provided through these filings is reliable and informative. The objective of this study is to evaluate the reliability and informativeness of one component of annual reports (Form 10-K): Item 1A. We will evaluate the informativeness of Item 1A in the setting of the bank industry during the financial crisis of 2007-2008.

Beginning in December 2005, the SEC required registrants to discuss ''the most significant factors that make the company risky'' under Risk Factors item (Item 1A) in quarterly and annual reports. The objective of Item 1A is to provide the securities market with timely information about potential future outcomes that may adversely affect the company's financial performance (SEC 2005). Item 1A being informative means that risks disclosed under this item should give necessary information about the risks that the company may face in the future to investors so they can make investment decisions accordingly. Besides being informative, risks disclosed under Item 1A should be timely. That is, they must be available to the decision makers before they lose capacity to influence decisions.[5] Risks listed under Item 1A may include but are not limited to market risk. Market risk or systematic risk are due to factors that affect the overall performance of the financial markets like recession, change in interest rate, political turmoil, etc.

However, any type of risk that is a threat to the future of the company should be disclosed under Item 1A such as specific risk. Specific risk unlike market risk is specific to a company.

Prior to mandating the inclusion of Item 1A in the quarterly and annual reports by the SEC, firms could voluntarily disclose their risk factors under the MD&A section of the annual report, and were only required to provide this information when making registration statements for equity and debt offerings on Form S-1 and foreign private issuers on Form 20-F. However, the registrants are now required to provide a full listing of risk factors in the 10-K, but they can skip filing the form 10-Q if they do not have new risks to disclose. The SEC's motivation for moving this requirement from the registration statements to the annual and quarterly reports was to provide investors with more timely and reliable information about registrant's changing risk environment.

The Risk Factors disclosed under Item 1A should "describe the most significant factors that may adversely affect the issuer's business, operations, industry or financial position, or its future financial performance" (SEC 2004). Although risk factors are required by the SEC in order to provide more timely information, managers might prefer to conceal bad news (Dye 2001) and this will cause some uncertainty about the informativeness of risk disclosures. Different studies on voluntary disclosure suggest that managers are more

[5] Kimmel, Paul D., Weygandt, Jerry J., and Kieso, Donald E. (2016) Financial Accounting Tools for Business Decision Making. Hoboken, NJ: Wiley.

willing to share a positive perspective of their firms, but they may be less transparent about bad news. Kothari et al. (2009) empirically support the theory that managers postpone the release of bad news to investors. This is also evident in Johnson (2010) which argues for a revision in regulations in order to increase the usefulness of risk disclosures. Although the SEC warned firms in 2010 to ‘‘avoid generic risk factor disclosure that could apply to any company’’, empirical studies suggest that the managers may have a tendency not to disclose specific risks because of career incentives (Watts and Zimmerman 1986; Fields et al. 2001). Because managers may have an incentive to conceal the true risks of their operations, the SEC mandated such disclosures as part of Item 1A.

The objective of this study is to evaluate the informativeness and timeliness of Item 1A in the setting of the bank industry during the financial crisis of 2007-2008. Banks should have anticipated the financial crisis and disclosed this in their risk disclosures. Banks were forefront of the crisis and they should have been the first to know about it. But it is possible that they only disclosed risks associated with the crisis after it began.

We develop Python scripts based on text mining techniques to be able to numerically represent the textual information. We count the number of risks and words disclosed in each Item 1A to observe the changes in the number of risks disclosed and the length of Item 1A. We also check the tone of disclosures using Loughran and McDonald (2011) word lists to identify the changes in the tone of the disclosures before and after the financial crisis. We also develop financial crisis dictionary (FCD) using two different methods to check the textual content of risk disclosures.

We find that the average number of risk factors disclosed by the national commercial banks has increased significantly from 2008 to 2009, compared to the increase from 2006 to 2007, indicating that on average they have started reporting more risks after the crisis. We qualitatively examine this issue by checking the underlying tone of these disclosures. The results indicate that the tone of Item 1A has become much more negative from 2008 to 2009. Our finings using the FCD suggest that appearance of the words describing the financial crisis have increased from 2008 to 2009. Overall, the results support the argument that the banking industry has failed to predict the financial crisis, and they have started noticing and disclosing risks related to the crisis only after it happened.

This paper contributes to the accounting literature by examining the effectiveness of Item 1A of the 10-K reports for the national commercial bank industry. This study analyzes the risk factors over a seven-year period. It also contributes to the literature through creation of a financial crisis-specific dictionary for text analytics research. Finally, this paper contributes by investigating the shareholders' perception of these risk factors during the crisis period around 2008.

The following sections are organized as follows: In section 2, we review the prior literature and background information on the financial crisis. Then we develop the hypotheses to test if the banking sector was able to predict the financial crisis. In section 3, we present the sample selection and data collection process. Research design is discussed in section 4, and results are presented in section 5. We conclude the study in section 6 and discuss some limitations.

2.2 Literature Review and Hypotheses Development

2.2.1 Risk Disclosure Literature Review

Early research in risk disclosure area shows a correlation between mandated quantitative disclosures and firms' market risk. Rajgopal (1999) provides evidence that commodity price risk disclosures required by the SEC are associated with stock market-based measures of commodity price risk exposure. Other studies such as Linsmeier et al. (2002) and Jorion (2002) prove a correlation between risk disclosures and equity prices. Prior literature has examined mandatory quantitative risk disclosures (Item 7A), and many empirical studies have looked at how investors use this quantitative information (Schrand 1997; Roulstone 1999; Wong 2001; Linsmeier et al. 2002). The concepts of risk and risk disclosure have received great attention in the literature in the recent years using textual analysis tools. Huang (2010) analyzes a small group of risk factors and develops a computer algorithm to identify risk factor headings and then uses key word analysis to determine whether one of his target 25 risk factors are identified in the 10-K. He finds mixed evidence that some key words are related to changes in risk and financial performance. Kravet and Muslu (2013) tests the relation between changes in companies' textual risk disclosures of 10-K filings and changes in stock market and analyst activity around the filings. Campbell et al. (2014) finds that firms that face greater risk disclose more risk factors, and market participants incorporate the information conveyed by risk factor disclosures into their assessments of firm risk and stock price, and that the disclosure decreases information asymmetry amongst firms' shareholders. Based on Campbell et al. (2014) findings investors try to use the information disclosed under risk factors to predict the riskiness of a firm and their stock price. Another study by Nelson and Pritchard (2014) investigates the shift of the risk disclosures from voluntary to mandatory regime. The results show that under the voluntary regime firms subject to greater litigation risk disclose more risk factors, update the language more frequently, and use more readable language compared to the firms facing lower litigation risk. Based on this research, these differences in the quality of disclosure are pronounced in the voluntary disclosure regime, but converge following the SEC mandate. Another study by Bao and Datta (2014) tries to quantify the different risk types from textual risk disclosures. They develop a variation of the latent Dirichlet allocation topic model and its learning algorithm for discovering and quantifying risk types from risk disclosures in 10-K. They find that around two-thirds of risk types lack informativeness and have no influence. Another study by Hope, Hu, and Lu (2014) propose a new measure called *Level of Detail*, which is computed based on an algorithm that quantifies the qualitative risk factor disclosures. Their results indicate that firms with longer and more readable 10-K report, lower accruals, greater risk, and smaller size tend to have more specific risk factor disclosures. They also find that the absolute value of the market reaction to the 10-K filing is positively and significantly associated with *Level of Detail*. This finding demonstrates that improved risk disclosures on 10-K enhance risk understanding and benefits the readers of the financial statement.

This study belongs to the research area of automated text analysis, which is intended for quantifying the underlying information in textual documents. O'Connor et al. (2011) argues that there is an increasing interest in the use of automated text analysis in the services of social science questions. They also argue that automated text analysis, which draws on techniques developed in natural language processing, information retrieval, text mining, and machine learning, should be properly understood as a class of quantitative social science methodologies. Bao and Datta (2014) classifies the methods for automated text analysis into three categories, including (1) dictionary, (2) supervised learning, and (3) unsupervised learning. They describe the dictionary method as a method which uses keywords or phrases to classify documents into categories or measure the extent to which documents belong to a particular category. They state that under supervised learning method the human coders first categorize a set of documents by hand, then they train a supervised model that automatically learns how to assign categories to documents using coded data. And they describe the unsupervised learning as a class of methods that learn underlying features of text without explicitly imposing categories of interests. Grimmer and Stewart (2013) state that the unsupervised learning methods are valuable since they could identify organizations of texts that are theoretically useful but perhaps understudied or previously unknown. For this study, we use a combination of dictionary and unsupervised learning methods which we describe in detail in the Research Design section.

2.2.2 Background Information on the Financial Crisis

In order to be able to analyze the changes in the Item 1A disclosures in regards to the financial crisis, a precise timeline of the beginning and the continuation of the crisis is needed. It is an impossible task to identify an exact starting point for the financial crisis.

Although there were early signs of distress even before 2007, a series of events in 2007 can be identified as early evidences. During 2007, the collapse of the housing bubble and the abrupt shutdown of the subprime lending led to losses for many financial institutions.

> Another evidence was the 1.5% drop in November 2006 of the ABX index.[6] In December, the same index fell another 3% after the mortgage companies Ownit Mortgage Solutions and Sebring Capital ceased operations. In January 2007, Mortgage Lenders Network announced it had stopped funding mortgages and accepting applications. In February, New Century reported bigger-than-expected mortgage credit losses and HSBC, the largest subprime lender in the United States, announced a $1.8 billion increase in its quarterly provision for losses. In March, Fremont stopped originating subprime loans after receiving a cease and desist order from the Federal Deposit Insurance Corporation. In April, New Century filed for bankruptcy. [7]

Also, in early 2007 the economy was beginning to show signs of stress with decline in home prices and oil prices above $75 a barrel.[7]

In late 2007, prices of AAA-rated private mortgage backed securities (PMBS) started to decline. A mortgage backed security is a type of asset-backed security that is backed by a mortgage or collection of mortgages. [8] one of the benefits of keeping a PMBS especially with AAA ratings, was that they were readily marketable. As housing prices began to fall in 2007, AAA-rated PMBS became unmarketable and they lost their value for liquidity purposes. Continuing in late 2007 and early 2008 banks suffered from billions of dollars in mortgage-related losses on loans, securities, and derivatives. Citigroup and Merril Lynch reported the biggest losses $23.8 billion and $24.7 billion respectively. [7]

The Bear Stearns collapse in 2008 was the first major casualty of the financial crisis.[9] The Bear Stearns companies, Inc. was a New York based investment bank and brokerage firm. After the fail of its hedge fund in July 2007, it faced more challenges in the second

[6] ABX index is a financial benchmark that measures the overall value of mortgages made to borrowers with subprime or weak credit. (http://www.investopedia.com/terms/a/abx-index.asp)

[7] Final Report of the National Commission on the Causes of the Financial and Economic Crisis in the United States. (January 2011)

[8] http://www.investopedia.com/terms/m/mbs.asp

[9] http://www.wsj.com/articles/SB124182740622102431

half of the year. The Federal Reserve Bank of New York provided an emergency loan to Bear Stearns in March 2008. However, the company could not be saved and was sold to JP Morgan chase (Ross 2008).

As the credit crisis started in August 2007 and with the failure of Bear Stearn hedge fund, Lehman's stock fell sharply, and finally on September 15, 2008, Lehman Brothers filed for bankruptcy (Malloy 2010). Lehman was the fourth largest investment bank in the United States and its collapse roiled global financial markets for weeks, given the size and status of the company as a major player in the U.S. and internationally. [10] To prevent further stress on global economy, the Federal Reserve provided an $85 billion two-year loan to AIG. [11] The series of events in September 2008 marked the worst financial disruption in postwar American history.[12]

Based on Final Report of the National Commission on the Causes of the Financial and Economic Crisis in the United States, banking supervisors failed to adequately and proactively identify and police the weaknesses of the banks and thrifts or their poor corporate governance and risk management, often maintaining satisfactory ratings on institutions until just before their collapse. Based on the timeline in Figure 2.1, several events occurred during 2007 and 2008 that can mark the beginning of financial crisis. Hence, we expect to see a pattern of disclosures describing the financial crisis starting on 2008 for the disclosures to be timely. And by 2009, disclosure of more risks describing the financial crisis is expected.

---

[10] http://www.investopedia.com/articles/economics/09/lehman-brothers-collapse.asp
[11] https://www.thebalance.com/aig-bailout-cost-timeline-bonuses-causes-effects-3305693
[12] Final Report of the National Commission on the Causes of the Financial and Economic Crisis in the United States. (January 2011)

2.2.3 Hypotheses Development

The SEC required registrants to disclose Risk Factors in quarterly and annual reports to provide the securities market with timely information about potential future outcomes that may adversely affect the company's financial performance. In order for these filings to provide timely and reliable information, the registrants have to be aware of the risks that a company may face even in extreme conditions. In this study, we evaluate the informativeness of Item 1A in the setting of the bank industry during the financial crisis of 2007-2008.

The focus of this study is to analyze if the risk disclosures of companies changed during the period of financial crisis. This study provides evidence as to whether the accounting system (1) provided evidence of greater risk ahead of the crisis; and (2) whether it reflected greater perceptions of risk during the crisis.

One factor in investigating whether banking section disclosed financial crisis-related risks before the event, is to look at the number of risks disclosed and the timing of the increase in the number of disclosed risks. As discussed before, there were a series of events and warning signs started on 2007 which led to the financial crisis. For the risk disclosures to be timely, the banking industry should have started adding risks that described the crisis ahead on their Item 1A on 2008. Alternatively, they may have wanted to hide or at least postpone the crisis that they were facing, so they may not have increased the number of risks disclosed to cover the risks related to financial crisis. We hypothesize that:

Hypothesis 1a: The number of risks disclosed by the national commercial banks increased before the financial crisis.

Hypothesis 1b: The number of risks disclosed by the national commercial banks increased after the financial crisis.

On of the ways of analyzing the content of Item 1A is to check the tone of the disclosures. Tone of the disclosure can be identified by looking at the frequency of negative words. If the banking industry had provided timely disclosures about the potential risk ahead, there should be an increase in the number of negative words and a more negative tone should be observed in 2008 filings. Therefore, we expect to see a more negative tone in the disclosures if the banking industry had warned the investors about the potential risks beforehand. But there is a chance that they started to disclose risks about the financial crisis only after the event, hence we see a more negative tone in the disclosures afterwards.

Hypothesis 2a: The tone of the risk disclosures (Item 1A) is more negative before the financial crisis.

Hypothesis 2b: The tone of the risk disclosures (Item 1A) is more negative after the financial crisis.

Besides checking the increase in the length of the Item 1A, checking the increase in the number of risks disclosed, and analyzing the tone of the disclosures we look at the content of this disclosures. By evaluating the content of this disclosures, we can check whether there were any risks describing the crisis ahead in the Item 1A of national commercial banks before the start of the financial crisis. Using text mining techniques to evaluate the content of Item 1A, we can identify whether banking industry have warned investor about financial crisis risk in advance. We expect the Item 1A disclosures to contain risks regarding the financial crisis beforehand. Alternatively, the banking industry may have postponed disclosing bad news and in that case the risks about the financial crisis appear in the disclosures afterwards.

Hypothesis 3a: Risks related to financial crisis were present in the Risk Factors section (Item 1A) of banks before the beginning of the financial crisis.

Hypothesis 3b: Risks related to financial crisis were present in the Risk Factors section (Item 1A) of banks after the beginning of the financial crisis.

The project will involve parsing the risk factors and their sub titles from 10-K forms of the national commercial banks (SIC: 6021) to test whether national commercial banks had foreseen the financial crisis in their annual and quarterly risk disclosures.

2.3 Data Collection Process

Since the interest of this study is to analyze the risk factors disclosed under Item 1A of national commercial banks, all available annual Form 10-K filings (in html format) are downloaded from the SEC's Electronic Data Gathering and Retrieval (EDGAR) system for the years 2006 through 2012. A total number of 62,410 10-Ks have been downloaded from EDGAR for this period with 18,419 unique CIKs. The CRSP/Compustat database lists both national commercial banks and state commercial banks under the SIC code 6020, and it does not differentiate between the two. However, Edgar uses SIC code 6021 specifically for national commercial banks. To be able to identify the 10-K files for national commercial banks from the ones downloaded from Edgar, first we match the CIK used as identifier in EDGAR with the CIK from CRSP/Compustat merged annual database. 2,627 10-K filings are matched with the 6020 SIC code in the CRSP/Compustst database. These 2,627 files include both national commercial banks and state commercial banks. To identify the 10-K files that specifically belong to national commercial banks, we check the CIK of these filings with EDGAR. 1,066 files out of 62,410 downloaded 10-K filings are for national commercial banks with the SIC code of 6021. Because some of these banks have gone through bankruptcy or merger, not all of them have 10-K files for all the years considered. All the banks that have gone through bankruptcy, or merger and acquisition have been dropped from the sample. We chose banks that had not gone through merger or defaulted to eliminate as many distracting and confounding factors as possible. On February 4, 2008 SEC exempted all the smaller reporting companies (public float of $75 million or less) from filling Item 1A. Based on this reform, smaller banks stopped disclosing any risks under this item. Only the banks that have reported Item 1A on their 10-K forms for every year of the period we are looking at are included in the sample. The final sample includes 128 national commercial banks.

Using a Python script, we extract the text between two titles of Item 1A and Item 2 (in some cases Item 1B) which is the following section in the 10-K filing. The text extracted is the content of Item 1A and is the risk factor disclosure needed for the analysis. The firms disclose their risks in the risk factors sections in the titles and descriptions format. Each title follows a description which explains the title of the risk in more detail. Using another Python script, we clean the text from all the html tags and divide the titles and bodies. So, for each Risk Factors item the "titles" and "descriptions" are separately written in text files.[13]

[13] See Appendix A for more details

Table 2.1 Sample selection

Total number of 10-K files downloaded from Edgar from Jan 2006 to Dec 2012	62,410
Merge with CRSP/Compustat merged to identify the files with SIC code 6020	2,627
Check with EDGAR to identify the files for national commercial banks (SIC code 6021)	1,066
Eliminating the banks which have gone through Bankruptcy, or merger and acquisition	948
Eliminating smaller reporting banks	896

2.4 Research Design

After downloading the 10-K filings and parsing the Item 1A, we develop a Python script which identifies each word and counts the number of times it appears in a document. Employing this program gives us the ability to identify the number of words each Item 1A contains. That number can be utilized to analyze the change in the length of this item over the desired time period.

Isolating the titles and descriptions for each bank from 2006 to 2012, we can identify the number of risks each bank has disclosed in each year of the time period that we are looking at. Knowing the number of risks that has been disclosed by each bank for each year, we can calculate the total and the average number of risks disclosed in each year, which will give us the ability to analyze the change in the risk disclosures quantitatively.

Previous studies in the automatic text retrieval literature indicate that term weighting approach will improve the effectiveness of an information retrieval system (Salton and Buckley, 1988; Buckley, 1993). As discussed in Salton and Buckley (1988) sets of single terms cannot provide complete identification of document content, and many enhancements in content analysis and text indexing procedures have been proposed since then to generate complex text representations more effectively. One of the developed methods is to employ word grouping methods of the kind provided by *thesauruses*, where classes of related words are grouped under common headings; the class headings can be used to more accurately analyze the information content of documents (Sparck Jones, 1971; Salton, 1972). We analyze the change in the tone of the Item 1A using this method and Loughran and McDonald (2011) Financial Sentiment Dictionaries. Loughran and McDonald Financial Sentiment Dictionaries (LMFSD) contains 3,532 unique words divided into six groups of positive, negative, uncertain, litigation, strong modal and weak modal.

Using LMSFD negative and positive dictionaries, we can count the frequency of the negative and positive words in the Item 1A, and by identifying a negativity score we can check how the tone of the documents change over the period. We identify two negativity scores to test for the tone of the disclosures. One negativity score is identified as the total number of negative words divided by the average number of words in Item 1A and the other one which is a stricter measure is identified as total number of negative words minus total number of positive words divided by average number of words in Item 1A. The second measure is a more strict one because we are deducting the number of positive words, and in many cases negation terms such as “not” occurs with the positive word.

The negativity scores will be identified as following:

$$Negativity\ Score_1 = \frac{Number\ of\ Negative\ Words}{Total\ Number\ of\ Words\ in\ item\ 1A}$$

$$Negativity\ Score_2 = \frac{Number\ of\ Negative\ Words - Number\ of\ Positive\ Words}{Total\ Number\ of\ Words\ in\ item\ 1A}$$

The scores are calculated for each year in the sample period.

$Number\ of\ Negative\ (Positive)\ Words$ is a word count of all the negative (positive) words; based on LMFSD negative (positive) words dictionary; for the Item 1A section of all the banks in our sample. $Average\ Number\ of\ Words\ in\ item\ 1A$ is the average number of words in each Item 1A for a specific year.

One way of evaluating the informativeness and timeliness of Item 1A in regards to disclosing the risks describing the financial crisis is to read through the documents to identify the risks that are describing the financial crisis and check the timing of appearance of them. This method requires a lot of time and effort, and is affected by human error. Our research methodology is to create a financial crisis-specific group of words, and check the frequency of those words in pre and post crisis years. To build a dictionary of financial crisis-specific words, we download 50 newspaper articles from Factiva with the financial crisis topic. After cleaning the articles which include stemming[14] and removing the stop words, we employ a Python script which identifies

[14] Stemming is the process of reducing the inflected words into their root or stem word. For example, “stems”, “stemming”, “stemmer”, and “stemmed”, all have similar meanings, and stemming technique will reduce them to “stem”.

each word and counts the number of times it appears in a document. This will help us identify the most frequent words of the articles. We pick the twenty most frequent words of the article, and those words form the financial crisis-related dictionary. After creating a dictionary related to the financial crisis, we look at the timing of appearing of these words in the filings. We check if the dictionary appears in the filings before or after the beginning of financial crisis.

Probabilistic topic models are a suite of algorithms whose aim is to discover the hidden thematic structure in large archives of documents (Blei; 2012). Latent Dirichlet Allocation (LDA) which is the simplest and most common topic model is a statistical model of document collections that tries to extract different topics from text (Blei, Ng, and Jordan; 2003). LDA determines the number of words in a document using the "bag of words", and makes the assumption that the order of the words in the document does not matter. Then, it groups words into desired number of topics, and finally assigns a probability to each word in each topic which shows the probability of that specific word appearing in the specific topic. Using a Python program, we apply the LDA model to the 50 chosen articles. We choose one as number of topics, because all of the articles are about financial crisis, and choosing more than one topic results in different categories of words that belong to the same topic. Employing the LDA model, and applying it to the articles results in a dictionary of the most probable words to appear in that topic with the probability of those words appearing in the topic. We check the frequency of the LDA dictionary's words in the Item 1A, and using the probability of the words in the dictionary as a weight of that word, we calculate a LDA score for each Item 1A. For example, "mortgage" and "subprime" appear in the topic with the probability of 0.013

and 0.010 respectively. The LDA score of a document with the word "mortgage" being repeated two times and the word "subprime" being repeated three times is 0.056. The Item 1A documents having higher LDA scores indicates the better description of financial crisis. Comparing the LDA scores between different years, we identify whether the banking section had started to describe the financial crisis before it starts, or they just started talking about the even after it happened.

2.5 Results

Table 2.2 shows a summary statistics of risk disclosures in Item 1A of the national commercial banks from 2006 to 2012. To identify total number of risks, we add up all the risks disclosed by national commercial banks for each year. Total number of risks has increased from 1,495 to 2,056 from 2008 to 2009, and from 1,409 to 1,495 from 2007 to 2008. The average number of risks is calculated by dividing the total number of risks by the number of banks in the sample for each year. The average number of risks disclosed has increased from 11.68 to 16.06 from 2008 to 2009, and from 11.01 to 11.68 from 2007 to 2008. There is a spike in the average number of risks reported in 2009 and 2010. The results indicate that they started to report more risks on average after the crisis. The spike in the total and average number of risks disclosed on 2009 indicate that the risk factor disclosures were not timely.

Total number of words in each Item 1A is calculated after stemming and eliminating the stop words, and is used to identify the average number of words per risk and percentage increase in the length of Item 1A.[15] The average number of words per risk is calculated by dividing the total number of words for all the Item 1A by total number of risks disclosed under Item 1A for each year. Average number of words per risk for all the banks in the sample has increased from 162.5 to 165.51 from 2007 to 2008. On the other hand, there is a significant increase in the number of words from 2008 to 2009, from 165.51 to 190.59. The number of words included in the Item 1A documents has increased by 58.37% from 2008 to 2009 which is much greater than the 8.07% increase from 2007 to 2008, indicating the fact that the Item 1A section disclosed by the national commercial banks has became lengthier after the financial crisis. Again, indicating that the banking sector did not make timely disclosures informing investors about the great crisis ahead.

Table 2.3 reports the number of positive and negative words, and both negativity scores for Item 1A for the banks in the sample. Number of positive and negative words are determined using Loughran and McDonald Financial Sentiment Dictionaries. The results indicate that the number of negative words has increased significantly from 2008 to 2009. For negativity $score_1$, we only look at the number of negative words because of the negation terms often used with a positive word. As a stricter test, we calculate negativity $score_2$ by subtracting the positive words from the negative words. For both negativity score, there is an increase from 2008 to 2009 showing that the Item 1A includes many more negative words in 2009. Negativity $score_1$ increases from 51.14 to 77.52 from 2008 to 2009. Negativity $score_2$ increases from 32.70 to 37.15 from 2007 to 2008, and from 37.15 to 60.11 from 2008 to 2009. The change in both negativity scores show that the

---

[15] Stemming is the process of reducing words to their root form. It is a popular technique in text mining research to reduce the number of inflected words.

national commercial banks start to use many more negative words in their Item 1A disclosure after the financial crisis.

To create the financial crisis-specific dictionary, 50 newspaper articles which discuss financial crisis are downloaded from Factiva. Figure 2.2 shows the distribution of the articles over time. Most of the articles are selected from 2008 and 2009 which is the time the financial crisis was discussed the most. Three articles are selected from 2007, 18 from 2008, 16 from 2009, seven from 2010, and six from 2011.

After cleaning the articles and running the Python program, we identify the top 20 most frequent financial words (stemmed) used in the articles. The list of 20 words form the financial crisis-specific dictionary.

Using the created dictionary, we check the frequency of the 20 words in the Item 1A disclosures. Table 2.5 reports the results, and it shows that the frequency of financial crisis dictionary increase

significantly in 2009 and 2010. We normalize the total number of times that the words of the dictionary appear in the Item 1A documents by dividing it by the number of risks disclosed in one year, and the total number of words used in Item 1A in that year. By dividing the frequency of the 20 words by total number of risks disclosed in that year, we observe that number of words describing the financial crisis increases from 9.96 to 10.91 in each risks from 2008 to 2009.

Figure 2.3 illustrates the spike from 2008 to 2009 more clearly. The results indicate that the national commercial banks in our sample start to discuss the risks related to financial crisis in their Item 1A after the even.

Applying the LDA model to the 50 articles which discuss the financial crisis results in a dictionary of the most probable words to appear in the topic with their probabilities. Table 6 shows the list of words generated by LDA with their probabilities.

Using the dictionary of financial crisis-related words that we create with LDA modal and using the probability of words as their weight, we create a score for each Item 1A based on the number of words used from the dictionary in the Item 1A disclosure. Adding up the scores, we assign a score to each year Item 1A disclosures. Normalizing the score by dividing it by number of risks disclosed each year and total number of words used in Item 1A sections in each year, we observe a more significant increase from 2008 to 2009 compared to 2007 to 2008.

Figure 2.4 illustrates the more significant increase in the LDA from 2008 to 2009 more clearly.

Figure 2.4 and the results of Table 2.8 indicate that banking sector start to disclose the risks related to financial crisis after it began.

2.6 Conclusion and Future Research

Beginning in 2005, SEC mandated all registrants to disclose the most significant factors that make the company risky under Item 1A. Investors consider the risks companies face to make financial decisions. In this study we investigate the timeliness and informativeness of risk factor disclosures for the banking industry. We examine the Item 1A of the national commercial banks from 2006 to 2012 to determine whether banking industry have warned the investors about risks related to financial crisis beforehand or they only disclose the risks related to financial crisis after the event. We evaluate the informativeness of Item 1A in regards to financial crisis both quantitatively and qualitatively. We find the number of risks disclosed increase after the financial crisis through counting all the risks disclosed by the national commercial banks in our sample. We see a spike from 2008 to 2009 in the number of risks disclosed which shows that national commercial banks started to disclose more risk after the crisis. We also investigate the risk disclosures qualitatively by checking the tone of this disclosures and their content. We find that the tone of the risk disclosures is more negative after the financial crisis. By creating a financial crisis-specific dictionary using the most common words in 50 articles which discuss the financial crisis and employing the LDA model, we check the timing of appearance of this dictionary in the Item 1A. We find that the warnings about financial crisis started to appear in the Item 1A section of the national commercial banks after the financial crisis, which shows the risk factors disclosed under Item 1A are not effective and timely in the financial crisis setting. This study has some limitations which open the possibility for future research. First, we dropped the bank that had gone through bankruptcy for our analysis to eliminate as many distracting and confounding factors as possible. But it would be interesting to look at those banks as a separate sample. Those banks were more affected by the financial crisis, thus looking at their risk disclosures will help to better analyze the risk factor disclosures. Also, bank holding companies were affected by the financial crisis as well. Broadening the sample of companies that we look at will help us to have better understanding and analysis of risk disclosures. Second, for this study we focus on the 10-K fillings, but the companies also disclose risks on their quarterly filings. Including the 10-Q fillings in the analysis of risk disclosures will give us a better vision about the exact timing of the disclosures.

Bibliography

Bao, Yang, and Anindya Datta. "Simultaneously discovering and quantifying risk types from textual risk disclosures." Management Science 60.6 (2014): 1371-1391.

Blei, David. 2012. Probabilistic topic models. Communications of the ACM 55(4):77–84.

Blei, David, Andrew Ng, and Michael Jordan. 2003. Latent dirichlet allocation. Journal of Machine Learning and Research 3:993–1022.

Buckley, C. (1993). The importance of proper weighting methods. In Proceedings of the workshop on Human Language Technology, pp. 349–352. Association for Computational Linguistics.

Campbell, J.L., Chen, H., Dhaliwal, D.S., Lu, H., Steele, L.B., 2014. The information content of mandatory risk factor disclosures in corporate filings. Review of Accounting Studies 19, 396–455.

Dye, R.A., 2001. An evaluation of “essays on disclosure” and the disclosure literature in accounting. Journal of Accounting and Economics 32, 181–235.

Fields, T.D., Lys, T.Z., Vincent, L., 2001. Empirical research on accounting choice. Journal of Accounting and Economics 31, 255–307.

Financial Crisis Inquiry Commission, and United States. Financial Crisis Inquiry Commission. The financial crisis inquiry report: Final report of the national commission on the causes of the financial and economic crisis in the United States. Public Affairs, 2011.

Grimmer J, Stewart BM (2013) Text as data: The promise and pitfalls of automatic content analysis methods for political texts. Political Anal. 21(3):267–297.

Hope, Ole-Kristian, Danqi Hu, and Hai Lu. "The benefits of specific risk-factor disclosures." Review of Accounting Studies, Forthcoming (2016).

Huang KW, Li ZL (2011) A multilabel text classification algorithm for labeling risk factors in SEC form 10-K. ACM Trans. Management Inform. Systems 2(3):1–19.

Johnson, S., 2010. SEC pushes companies for more risk information. CFO Magazine.

Jorion, P. 2002. How Informative Are Value-at-Risk Disclosures? The Accounting Review 77: 911-931.

Kothari, S., S. Shu, and P. Wysocki, 2009. Do managers withhold bad news? Journal of Accounting Research 47: 241-276.

Kravet, T., Muslu, V., 2013. Textual risk disclosures and investors' risk perceptions. Rev Account Stud 18, 1088–1122.

Linsmeier, T.J., Thornton, D.B., Venkatachalam, M., Welker, M., 2002. The Effect of Mandated Market Risk Disclosures on Trading Volume Sensitivity to Interest Rate, Exchange Rate, and Commodity Price Movements. The Accounting Review 77, 343–377.

Loughran, T. and B. McDonald (2011). When is a Liability not a Liability? Textual Analysis, Dictionaries, and 10-Ks. Journal of Finance 66(1), 35 – 65.

Malloy, Michael P. (2010). Anatomy of a Meltdown: A Dual Financial Biography of the Subprime Mortgage Crisis. Wolters Kluwer Law & Business. ISBN 978-0-7355-9458-6.

Nelson, Karen K. and Pritchard, Adam C., "Carrot or Stick? The Shift from Voluntary to Mandatory Disclosure of Risk Factors" (2014). Law & Economics Working Papers.

O'Connor B, Bamman D, Smith NA (2011) Computational text analysis for social science: Model assumptions and complex- ity. Proc. 2nd Workshop Comput. Soc. Sci. Accessed March 22, 2014, http://people.cs.umass.edu/wallach/workshops/nips2011css/papers/OConnor.pdf.

Rajgopal, S. 1999. Early Evidence on the Informativeness of the SEC's Market Risk Disclosures: The Case of Commodity Price Risk Exposure of Oil and Gas

Producers. The Accounting Review 74: 251-280.

Ross, Andrew (March 17, 2008). "JP Morgan Pays $2 a Share for Bear Stearns". The New York Times. Retrieved on September 30, 2008.

Roulstone, D.T., 1999. Effect of SEC Financial Reporting Release No. 48 on Derivative and Market Risk Disclosures. Accounting Horizons 13, 343–363.

Salton, G. A new comparison between conventional indexing (Medlars) and automatic text processing (Smart). J. ASIS. 23, 2, (March-April 1972), 75-84.

Salton, G. and C. Buckley (1988). Term-weighting approaches in automatic text retrieval. Information Processing & Management 24(5), 513 – 523.

Securities and Exchange Commission (SEC). 2004. Securities Offering Reform. Release Nos. 33-8501, 34- 50624. Available at: http://www.sec.gov/rules/proposed/33-8501.htm

Securities and Exchange Commission (SEC). 2005. Securities Offering Reform. Release Nos. 33-8591, 34- 52056. Available at: http://www.sec.gov/rules/final/33-8591.pdf

Securities and Exchange Commission (SEC). 2007. Smaller Reporting Company Regulatory Relief and Simplification. Release Nos. 33-8876, 34-56994:39-2451. Available at: http://www.sec.gov/rules/ final/2007/33-8876.pdf

Schrand, C.M., 1997. The Association between Stock-Price Interest Rate Sensitivity and Disclosures about Derivative Instruments. The Accounting Review 72, 87–109.

S. K. Jones, “A statistical interpretation of term specificity and its applications in retrieval,” Journal of Documentation, vol. 28, no. 1, pp. 11–21, 1972.

Watts, R.L., Zimmerman, J.L., 1986. Positive accounting theory. Englewood Cliffs, NJ: Prentice Hall.

Wong, M.H.F., 2000. The Association between SFAS No. 119 Derivatives Disclosures and the Foreign Exchange Risk Exposure of Manufacturing Firms. Journal of Accounting Research 38, 387–417.

2.7 Tables for Chapter 2

Table 2.2 Summary Statistics of Risk Disclosures

128 Banks in the Sample	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**
Total Number of Risks Disclosed	1,287	1,409	1,495	2,056	2,326	2,344	2,516
Average Number of Risks Disclosed	10.05	11.01	11.68	16.06	18.17	18.31	19.66
Min Number of Risks Disclosed	4	4	4	4	4	6	9
Max Number of Risks Disclosed	26	27	32	41	54	57	56
Std.	4.82	5.00	5.48	7.15	9.54	10.29	10.49
Average Number of Words Per Risk	160.33	162.50	165.51	190.59	200.52	207.21	214.34
Percentage Increase in Length of Item 1A		10.96%	8.07%	58.37%	19.03%	4.13%	11.03%

This table shows a summary statistics of risk disclosures. Total number of risks disclosed by all the companies in the sample for each year, and average number of risks disclosed for each year are reported in this table. We also report the increase in the length of Item 1A on this table. Based on this results, Item 1A increases 58.37% from 2008 to 2009.

Table 2.3 Number of Positive and Negative Words

128 Banks in the Sample	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**
Number of Positive Words in All Item 1A	2,047	2,214	2,316	3,317	3,630	3,809	4,211
Avg. Number of Positive Words Per Risk	1.59	1.57	1.55	1.61	1.56	1.63	1.67
Min Number of Positive Words Per Risk	0	0	0	0	0	0	0
Max Number of Positive Words Per Risk	19	20	20	18	18	18	21
Std.	2.39	2.38	2.35	2.32	2.20	2.67	2.41
Number of Negative Words in All Item 1A	6,575	7,528	8,464	14,774	17,289	17,564	20,128
Avg. Number of Negative Words Per Risk	5.11	5.34	5.66	7.19	7.43	7.49	8.00
Min Number of Negative Words Per Risk	0	0	0	0	0	0	0
Max Number of Negative Words Per Risk	50	54	66	64	60	83	60
Std.	5.19	5.26	5.85	7.40	7.33	7.60	8.28
Negativity Score 1	41.01	46.33	51.14	77.52	86.22	84.77	93.91
Negativity Score 2	28.24	32.70	37.15	60.11	68.12	66.38	74.26

Table 2.4 List of 20 Most Frequent Stemmed Words of the Articles

Subprime	Rate
Mortgage	Borrow
Loan	Price
Credit	Financ
Hous	Interest
Default	Risk
Market	Grow
Percent	Foreclos
Capital	Crisis
Los	Lend

To create a financial crisis dictionary, we download 50 articles which describe the financial crisis of 2007-2008. We pick a list of the 20 most frequent words based on these articles to be the financial crisis dictionary.

Table 2.5 Frequency of Financial Crisis Dictionary in the Item 1A

128 Banks in the Sample	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**
Frequency of FCD in Item 1A	12,250	13,587	14,893	22,437	25,545	25,959	28,306
Min Number of FCD Per Risk	0	0	0	0	0	0	0
Max Number of FCD Per Risk	72	76	76	81	74	98	96
Std.	10.67	10.67	11.29	11.92	11.71	11.89	11.98
Divided by Number of Risks	9.518	9.643	9.962	10.913	10.982	11.075	11.250
Divided by Avg. Number of words	76.403	83.612	89.983	117.722	127.392	125.281	132.064
Divided By Number of Words	0.059	0.059	0.060	0.057	0.055	0.053	0.052

This table reports the frequency of the financial crisis dictionary (FCD) in the Item 1A of all the companies in our sample for each year. The number of FCD words increase in Item 1A from 14,893 to 22,437 from 2008 to 2009 which shows that the banking industry started to disclose more risks about the financial crisis after it began.

Table 2.6 Dictionary of Financial Crisis Words Created using LDA Model

Probability	**Word**	**Probability**	**Word**
0.013	Mortgage	0.006	Market
0.012	Rate	0.005	Credit
0.01	Subprime	0.005	Financ
0.009	Loan	0.005	Percent
0.008	Hous	0.005	Origin
0.007	Bank	0.004	Foreclos
0.007	Price	0.004	Risk
0.006	Default	0.004	Interest
0.006	Borrow	0.003	Repa

This table reports the words created by LDA which is a probabilistic topic model. LDA assigns a probability to each word in each topic which shows the probability of that specific word appearing in the specific topic. We create a second financial crisis dictionary using this method.

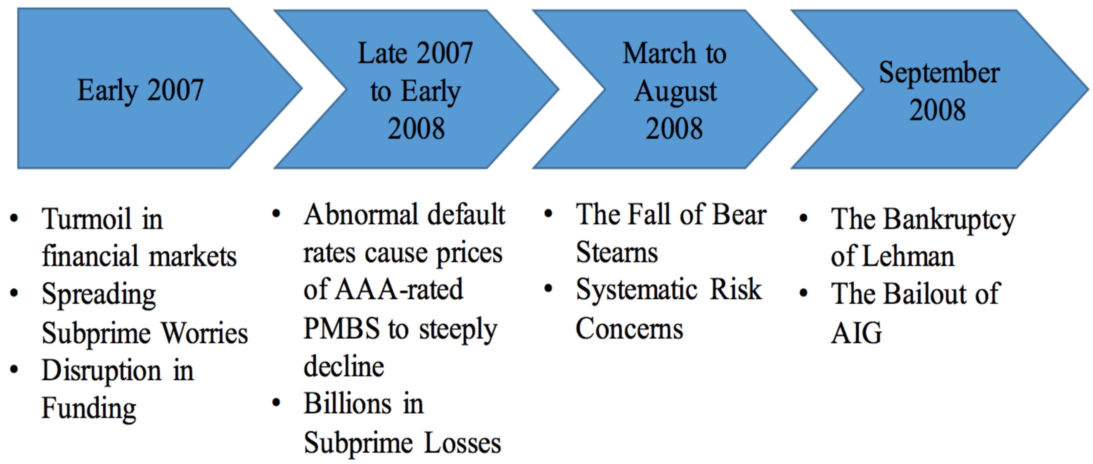
Table 2.7 Frequency of Financial Crisis Dictionary Created Using LDA in the Item 1A

128 Banks in the Sample	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**
Score of Item 1A based on LDA FCD	64.9	71.61	79.09	117.84	133.27	135.54	147.4
Lowest Score	0	0	0	0	0	0	0
Highest Score	0.465	0.489	0.533	0.478	0.0534	0.0534	0.0596
Std.	0.062	0.062	0.066	0.067	0.065	0.066	0.067
Divided by Number of Risks	0.05043	0.05082	0.05290	0.05732	0.05730	0.05782	0.05859
Divided by Avg. Number of words	0.40479	0.44068	0.47786	0.61829	0.66462	0.65412	0.68769
Divided By Number of Words	0.00031	0.00031	0.00032	0.00030	0.00029	0.00028	0.00027

This table reports the results generated by LDA financial crisis dictionary. We check the frequency of the LDA dictionary's words in the Item 1A, and using the probability of the words in the dictionary as a weight of that word, we calculate a LDA score for each Item 1A. For example, "mortgage" and "subprime" appear in the topic with the probability of 0.013 and 0.010 respectively. The LDA score of a document with the word "mortgage" being repeated two times and the word "subprime" being repeated three times is 0.056. The Item 1A documents having higher LDA scores indicates the better description of financial crisis. The results are similar to the previous test. We see an increase in using the LDA FCD from 2008 to 2009.

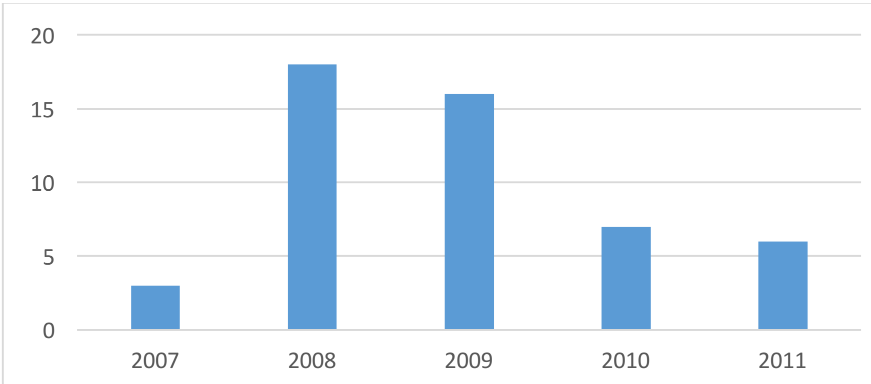
2.8 Figures for Chapter 2

Figure 2.1 A Timeline for Financial Crisis of 2007-2008



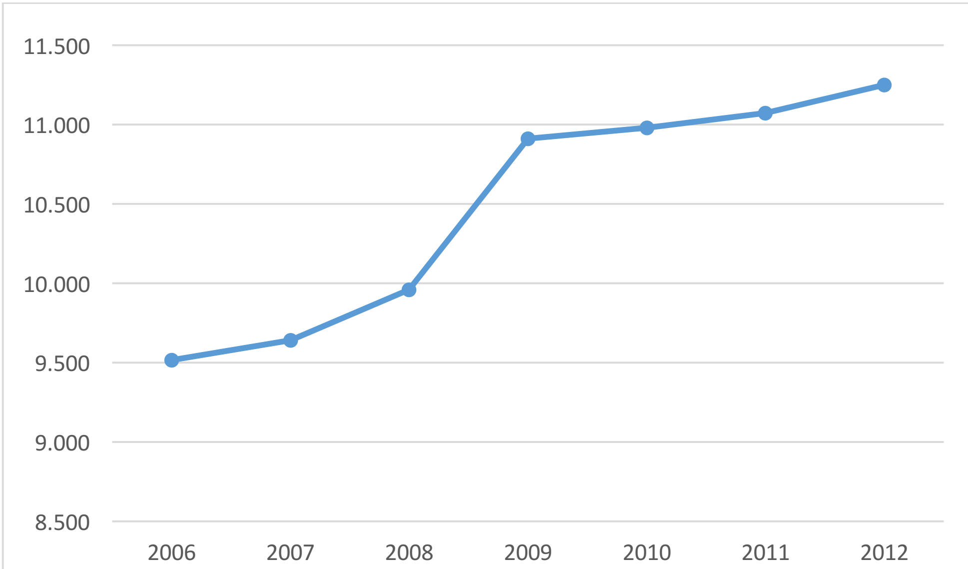
This figure shows a timeline of financial crisis of 2007-2008. As we see in this figure, there were early signs of financial crisis even on early 2007. Several events occurred during 2007 and 2008 that can mark the beginning of financial crisis. Hence, we expect to see a pattern of disclosures describing the financial crisis starting on 2008 for the disclosures to be timely. And by 2009, disclosure of more risks describing the financial crisis is expected.

Figure 2.2 Distribution of the 50 Articles over Years



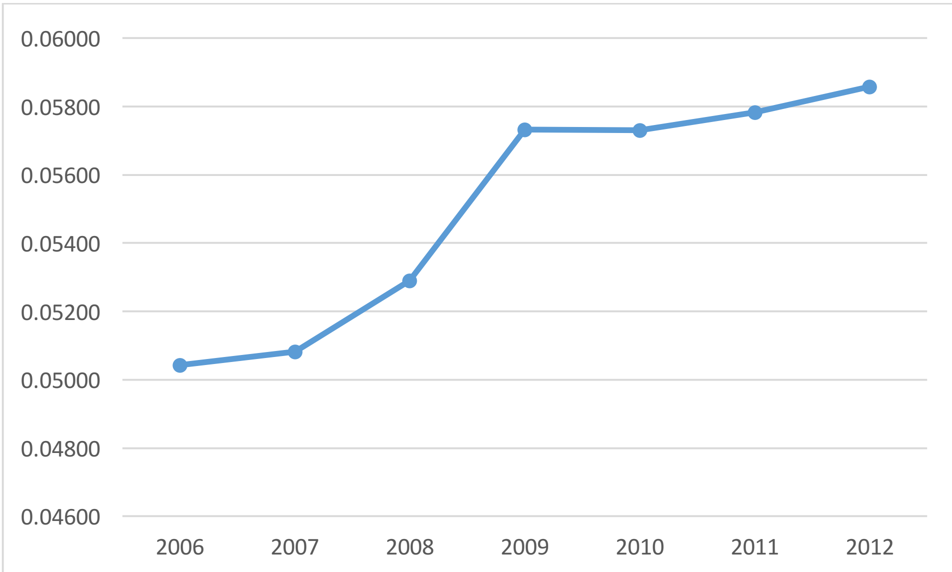
To create the financial crisis-specific dictionary, 50 newspaper articles which discuss financial crisis are downloaded from Factiva. This figure shows the distribution of the articles over time. Most of the articles are selected from 2008 and 2009 which is the time the financial crisis was discussed the most. Three articles are selected from 2007, 18 from 2008, 16 from 2009, seven from 2010, and six from 2011.

Figure 2.3 Frequency of the Financial Crisis Dictionary in the Item 1A



This figure is based on Table 2.5. The banking industry started to disclose more risks about financial crisis from 2008 to 2009 and the spike is clearly indicating that they started talking about the financial crisis after it began.

Figure 2.4 Change in LDA Score



This figure is based on the LDA FCD results. We find that based on the LDA model the banking company started to use more words which describe financial crisis from 2008 to 2009. And the spike we see in this figure proves that the only started to disclose risks about financial crisis after in began.

2.9 Appendix for Chapter 2

2.9.1 Appendix A

After extracting the Risk Factors item, an “html to text” converter program is used to convert the html files to plain text. As mentioned before the format of the disclosures are titles and descriptions. Each company discloses some “risk titles” and “description” for them. The titles mostly appear in bold or italic formats (html tags for bold could be <b>, <bold>, or <strong> and for italics could be <italics> or <i>). The “html to text” program would also mark up the bold and italic titles with desired symbols; in this case I use plus signs for a bold title and underline signs for italics titles. Using this program, the files are marked up with pluses and underlines which make it easier to isolate the titles and descriptions.