The Effect of Optionability on Underlying Stock Prices

A thesis submitted in partial fulfilment of the

requirements for the degree of

Master of Commerce in Finance

University of Canterbury

Christchurch, New Zealand

By Øyvinn Døhl Rimer

2006

Senior Supervisor: Prof. Edwin Maberly

Co-Supervisor: Dr. Raylene Pierce

Abstract

In Ni, Pearson and Poteshmans' (2005) *Journal of Financial Economics*-article, they claim that the expiration-day price-distribution of optionable stocks is subject to inefficiencies caused by stock price manipulation and portfolio rebalancing by delta hedgers. In this thesis, two main shortcomings of Ni et al.'s (2005) study are identified. In particular, they appear to have been ignorant of fundamental microstructure factors, and they did not derive an expression to represent the theoretical price-distribution of the relevant assets. After accounting for essential microstructure variables, and calculating the theoretical distribution, results that contradict Ni et al. (2005) are found. In particular, optionable stocks are found to experience efficiency gains on expiration days, and the distribution of underlying asset prices is closer to its theoretical benchmark on expiration days relative to non-expiration days.

Acknowledgements

I would like to express my sincere gratitude to my chief supervisor, Professor Ed Maberly, for being an excellent guide, supporter, critic and teacher throughout this research. His passion for Finance is enormous and without his enthusiasm, this research would have been unachievable. I would also like to thank Dr. Raylene Pierce, who has provided invaluable comments, motivation and support throughout this research. Her comments on the structure of this research have made this thesis significantly more readable.

I would like to extend my thanks to Dr. Philip Meguire and Dr. John Fountain for helping me to identify appropriate test statistics and for introducing me to relative entropy. Furthermore, I am thankful to Dr. Philip Gunby and Dr. Karen Lewis for providing extensive and insightful comments on the research proposal.

I would like to thank my parents for always being positive and supportive of my choices in life. Their friendship, love and support, have given me strength and confidence to reach any goal of mine.

At last, extraordinary appreciation goes out to Esther Song, who has been my rock throughout my studies. She has taught me that hard work can take a person wherever they want to go. She has provided companionship, encouragement, friendship and love and if it was not for her, I would not have made it this far.

Table of Content

ABSTRACT II
ACKNOWLEDGEMENTS III
TABLE OF CONTENT IV
TABLE OF FIGURES V
LIST OF TABLES VI

CHAPTER 1: INTRODUCTION 1
1.0 SIGNIFICANCE OF THE RESEARCH 1
1.0.1 Researchable Questions 3
1.1 ORGANISATION OF THESIS 3

CHAPTER 2: CONCEPTS AND LITERATURE 6
2.0 CONCEPT: OPTIONS 6
2.1 LITERATURE 7
2.1.1 The Efficient Markets Hypothesis 7
2.1.2 Behavioural Finance and Financial Anomalies 9
2.1.3 Market Microstructure 12
2.1.4 Financial Derivatives 14
2.1.4.1 Expiration-day effects 15

CHAPTER 3: METHODOLOGY 18
3.0 DATA COLLECTION AND THE INTEGRITY OF SOCIAL SCIENCE DATA 18
3.1 TESTS AND ROBUSTNESS 22

CHAPTER 4.0: RESULTS 25
QUESTION 1: ARE THERE MARKET MICROSTRUCTURE DIFFERENCES BETWEEN OTC, AMEX AND NYSE STOCKS; IF THERE ARE, HOW DO THEY INFLUENCE RESEARCH OF STOCK PRICES BEHAVIOUR ON OPTION EXPIRATION DAYS? 25
4.1 Automatic exercise 26
4.2 Strike price increments 29
4.3 The underlying assets 32
4.3.1 Stealth Options 33
4.4 Summary 35
QUESTION 2: WHAT DISTRIBUTION CONSTITUTES THE EXPECTED CLUSTERING OF STOCK PRICES? 36
4.5 The evidence 38
4.5.1 Wider intervals 40
4.6 Optionability and the uniform distribution 43
4.6.1 Decimal intervals 44
4.7 The S&P 600 49
4.7.1 Larger intervals 49
4.8 Over the counter vs. exchange traded stocks 51
4.8.1 Order preferencing 51

4.8.2 The number of market makers *52*
4.9 Clustering of S&P 600 Stocks on 25-Cent intervals *53*
A. Nasdaq *53*
B. Listed Stocks: NYSE and AMEX *54*
4.10 Summary *56*
QUESTION 3: IS THE EFFICIENCY OF OPTIONABLE AND NON-OPTIONABLE STOCKS DIFFERENT; IF IT IS, IS THE DIFFERENCE MORE PROMINENT ON OPTION EXPIRATION DAYS? 57
4.11 Optionable vs. non-optionable stocks. *59*
4.12 Results *60*
4.13 The S&P 600: A comprehensive analysis *62*
4.13.1 Optionable stocks *62*
4.13.2 Non-optionable stocks *64*
4.14 Triple witching *66*
4.15 Regression analysis *68*
4.15.1 Logistic Regression *69*
4.15.2 OLS Regression *71*
4.16 Summary *74*
CHAPTER 5.0: CONCLUSIONS 75
5.1 WAS THE OBJECTIVE ACHIEVED? 76
5.2 LIMITATIONS AND KEY ASSUMPTIONS 76
5.3 FURTHER RESEARCH 77
REFERENCES 78
THESIS APPENDICES 82
APPENDIX 1: CLUSTERING ON OPTIONS STRIKE PRICES 82
A1: Optionable stocks *82*
A2: Non-optionable stocks *84*
APPENDIX 2: TRIPLE WITCHING 86
APPENDIX 3: LOGISTIC REGRESSION 87
APPENDIX 4: S&P 600: CLUSTERING ACROSS DECIMALS 88
APPENDIX 5: AN EXTENSION OF THE ANALYSIS IN SECTION 4.9 90
APPENDIX 6: STATA CODE – DO-FILE 91

Table of Figures

Figure 1: Strike Deltas and Expected Clustering 31
Figure 2: Observed clustering across all stocks – decimals 39
Figure 3: The Cumulative Distribution Function 40
Figure 4: Clustering across all stocks – 25-cent intervals. 41
Figure 5: S&P 600 clustering of OS and NOS – decimals 47
Figure 6: S&P 100 and Nasdaq 100 clustering – decimals. 48

Figure 7: S&P 600 clustering; OS vs NOS – 25-cent intervals. 50
Figure 8: S&P 600, Nasdaq only, OS vs NOS – 25-cent intervals..... 54
Figure 9: S&P 600, Listed only, OS vs NOS – 25-cent intervals..... 55
Figure 1A: S&P 600 clustering on $[K \pm 24\,cents]$ – Listed – OS 82
Figure 1B: S&P 600 clustering on $[K \pm 14\,cents]$ – Listed – OS..... 82
Figure 1C:. S&P 600 clustering on $[K \pm 24\,cents]$ – Nasdaq – OS..... 83
Figure 1D: S&P 600 clustering on $[K \pm 14\,cents]$ – Nasdaq – OS. 83
Figure 1E: S&P 600 clustering on $[K \pm 24\,cents]$ – Listed – NOS. 84
Figure 1F: S&P 600 clustering on $[K \pm 14\,cents]$ – Listed – NOS 84
Figure 1G: S&P 600 clustering on $[K \pm 24\,cents]$ – Nasdaq – NOS 85
Figure 1H: S&P 600 clustering on $[K \pm 14\,cents]$ – Nasdaq – NOS. 85
Figure 2A: Triple witching clustering on $[K \pm 24\,cents]$ Listed & Nasdaq; OS & NOS..... 86
Figure 2B: Triple witching clustering on $[K \pm 14\,cents]$ Listed & Nasdaq; OS & NOS..... 86
Figure 4A: OS, non-expirations – decimals. 88
Figure 4B: OS, expirations – decimals..... 88
Figure 4C: NOS, non-expirations – decimals..... 89
Figure 4D: NOS, expirations – decimals..... 89
Figure 5A: Expiration vs non-expiration; Listed vs Nasdaq – 25-cent intervals..... 90

List of Tables

Table 1: Moneyness and automatic exercise 27
Table 2: Clustering across 25-cent intervals – all stocks..... 46
Table 3: Clustering test statistics: S&P 100, Nasdaq 100, and S&P 600 61
Table 4: S&P 600 clustering before, at, and after expiration Fridays – OS 63
Table 5: S&P 600 clustering before, at ,and after expiration Fridays – NOS 65
Table 6: S&P 600 Triple witching clustering..... 67
Table 7: Marginal effects of regressions 71

Chapter 1: Introduction

The objective of this thesis is to study the impact of option expiration days on the options' underlying assets. Ni, Pearson and Poteshman (2005) identify a strange phenomenon that occurs amid optionable stocks on option expiration days; namely, they find that optionable stock prices are pulled (or pushed) towards options strike prices. They conclude that the alteration of optionable stocks is induced by market manipulation by firm proprietary traders, with net written options positions, who try to force their written option positions from expiring In-The-Money (ITM).[1] This translates to powerful investors being able to manipulate the price of underlying assets for financial gain, at the cost of uninformed, unsophisticated investors. Certainly, if these allegations are true, industry regulators and policymakers need to initiate programmes that deal with the problems quickly; whether solutions involve tighter overseeing of the marketplace, suspensions, fines, or jail-terms, something has to be done. However, this thesis identifies a number of inaccuracies in Ni et al.'s (2005) analysis, such as incorrect strike price increments, the absence of a theoretical clustering, and insensitivity to microstructure-features. When these inaccuracies are accounted for, the inefficiencies of optionable stocks, as described by Ni et al. (2005), largely disappear. In fact, given that the assumptions made in this study are stable and acceptable, it is clear that optionable stocks experience an efficiency gain on expiration days.

1.0 Significance of the research

As mentioned above, there is one chief research objective; namely, to extend the literature on the behaviour of optionable stock prices on option expiration days. Stock

[1] The move towards options strike prices is referred to as *clustering*, which is the terminology used in Ni et al.'s article and for the remainder of this thesis.

price clustering of optionable stocks, at or near options strike prices is, according to Ni et al. (2005), abnormally high on expiration Fridays. Their finding is a major motivator of this research because it contradicts established theory, which states that optionability, or derivative trading, increases informational efficiency. In order to test the efficiency of the market for optionable stocks, it is important to understand the market microstructure of the relevant markets. Furthermore, a benchmark, with which the observed behaviour of stock prices can be compared, is required. Therefore, to facilitate the tests of the main research objective, two preliminary analyses are required. First, it is necessary to investigate the underlying microstructure of the assets being researched. In particular, stocks listed on the New York Stock Exchange (NYSE) and the American Stock Exchange (AMEX) have fundamentally different market structures relative to stocks traded on the National Association of Securities Dealers Automated Quotation system (Nasdaq). Therefore, sensitivity to different characteristics of these markets is important to the subsequent clustering analysis. Other essential microstructure features of this thesis include automatic exercise thresholds, strike price increments, the nature of the underlying asset, and stealth clustering. Second, an expression that represents the expected clustering of stock prices has to be derived. The supporting paradigm of this thesis is that the distribution of stock prices ought to be uniformly distributed across given intervals. The uniformity assumption drives the majority of tests in this thesis.

1.0.1 Researchable Questions

It is important to note that for each general research question, one or more specific questions require an answer. These are discussed and tested within the sections in which they are required. However, the general researchable questions can be summarised as:

1. Are there market microstructure differences between OTC, AMEX and NYSE stocks; if there are, how do they influence research of stock price behaviour on options expiration days?

2. What distribution constitutes the expected clustering of stock prices?

3. Is the pricing efficiency of Optionable and Non-Optionable stocks different, but particular on expiration days?

1.1 Organisation of thesis

The structure of this thesis is as follows. Chapter 2 provides a basic introduction of options and the way that they provide payoffs for their owners; a thorough discussion of relevant existing knowledge, which includes the Efficient Markets Hypothesis literature and its rival, the Behavioural Finance literature; a brief review of major studies in market microstructure; and research of financial derivatives and their impact upon the underlying assets.

Chapter 3 focuses on data collection, data accuracy and the statistical tests employed to discriminate between expiration day clustering and non-expiration day clustering.

In Chapter 4, the results of the three main research questions are presented. The three questions are answered consecutively and Questions One focuses on the importance of market microstructure for empirical research in Finance. In particular, differences in automatic exercise rules among customer accounts, firm- and market maker accounts are highlighted. Furthermore, different strike price increments (strike deltas) are important for the calculation of the expected clustering; therefore, a discussion of strike price increments illustrates how they vary between stocks. The characteristics of the underlying asset are also considered, because, for example, the laws of arbitrage govern assets such as Exchange Traded Funds, which makes it practically impossible to manipulate the underlying asset. Hence, it is important that only assets that, in theory, can be manipulated are included in the analysis. Awareness of the relevant microstructure is material to the results of the clustering analysis and Chapter 4 introduces the main microstructure variables of interest.

Question 2 focuses on the distribution of asset prices of Optionable and Non-Optionable stocks. Exciting findings of efficiency gains from options trading are presented. In this section, a major analysis of the distributions of optionable stocks, non-optionable stocks, listed stocks, and Nasdaq stocks fulfils the objective of establishing whether it is valid to assume that the distribution of stock prices is uniform. In particular, uniformity across 25-cent intervals is consistent with uniformity across the automatic exercise thresholds. Uniformity on these intervals permits the use of the uniform distribution as the main paradigm and as the benchmark of efficiency for stock price distributions. Furthermore, order preferencing rules and the number of market makers following each stock may have material impacts upon the uniformity of the stock price distribution.

The ultimate thesis question, Question Three, is answered at the end of Chapter 4. Here, clustering on the $[K \pm 14\,cents]$ and $[K \pm 24\,cents]$ − intervals, of optionable and non-optionable stocks, is split into three periods: the second, third and fourth Fridays of every month.[2] This allows for comparisons of clustering on expiration days as opposed to non-expiration days. The sample is further spilt into appropriate sub-samples to account for listed versus Over-The-Counter (OTC) stocks. In addition, a Logistic regression analysis is employed to study the impact that options data, such as open interest and volume have on clustering. The results of this section are fascinating because optionability appears to have positive effects upon efficiency. No evidence that optionable stocks are being manipulated is identified, which contradict the results of Ni et al. (2005).

Chapter 5 summarises main findings and provides conclusions. Further research, which is likely to lead to an enhanced understanding of options markets and to the behaviour of optionable stocks, is suggested. Following Chapter 5 are the references and thesis appendices.

[2] The third Friday of every month corresponds to option expiration (i.e. "expiration Friday").

Chapter 2: Concepts and Literature

2.0 Concept: options

An option is a financial product and the price of an option is derived from an underlying asset. There are two types of options: *call* options and *put* options. A call option gives the owner the right, but not the obligation, to purchase the underlying asset at a pre-specified price. A "real-world" example of a product that closely resembles a call option is a pizza voucher. A pizza voucher gives the owner of the voucher the right, but not the obligation, to purchase the *underlying asset*, in this case a pizza, at the price that is given on the voucher. A voucher that gives the owner the right to purchase a pizza for $10 would benefit the owner if a pizza, without the voucher, costs more than $10.[3] If the price for the pizza, without the voucher, is $15, then the owner of the voucher receives a $5 payoff by redeeming (exercising) the voucher. In the world of finance, options can be traded on underlying assets such as stocks, index funds, commodities, interest rates and exchange rates. The pre-agreed purchasing price is referred to as the option's *strike price*. In the pizza example, the strike price is $10. In this thesis, the underlying assets are stocks. Assuming that a stock that can be purchased in the market for $20 has an exchange traded call option with a strike price of $15. The owner of such an option has the right to purchase the underlying stock at a price of $15. It is obvious that being able to buy a stock at $15 and resell it in the market for $20 would produce a payoff of $5, ignoring transaction costs.

A put option gives the owner of the option the right, but not the obligation to sell the underlying stock in the market at a pre-specified price. Hence, if an investor expects

[3] It is assumed that the owner of the voucher wants to buy the pizza in the first place.

that a particular stock's price will decrease in the future, then they can purchase a put option that gives them the right to sell the stock. For example, if the current stock price is $20 and they believe that the price will drop to $15, then they can buy a put with a strike price of $20, which allows them to sell the stock at $20. Therefore, they can purchase the stock for $15, sell it again for $20 and receive an instant payoff of $5, ignoring transaction costs.

2.1 Literature

The relevant literature for this study covers four broad groups: the Efficient Markets Hypothesis, Behavioural Finance and Financial Anomalies, Market Microstructure and Financial Derivatives. Contained within these broad groups are several sub-groups that deserve particular attention for the subsequent analyses.

2.1.1 The Efficient Markets Hypothesis

The Efficient Markets Hypothesis (EMH) emerged out of the University of Chicago in the 1960s (Shleifer, 2002). Some of the founding studies of this theory are by Fama (1965 and 1970), Granger and Morgenstern (1970), and Sharpe (1964). Much of the early research of the EMH focuses on the unpredictability of asset returns, in particular that asset prices follow a random walk.

The ultimate message from the vast EMH literature is that abnormal, risk adjusted returns are unattainable given access only to public information (Summers, 1986). In other words, the EMH is that investors cannot earn above average returns without taking on above average risks (Malkiel, 2003). There are three forms of market efficiency: weak form, semi-strong form and strong form. All three forms are consistent with the above statement that investors are virtually unable to earn abnormal returns on investments after appropriately adjusting for the risks of their

investments. Through the 1980s, the EMH faced enormous scrutiny and criticism from an alternative school-of-thought; namely, the behavioural finance school-of-thought, which argues that, owing to certain investor behaviours, informed investors can profit from the anomalous behaviour of particular groups of investors. A discussion of the behavioural literature is provided below. Throughout the 1990s, however, there was a large rise in articles supporting the EMH. The renaissance of the EMH is largely driven by researchers' increased understanding of the models employed by the behavioural literature. EMH-followers have learned that a majority of the models and assumptions that are utilised in the behavioural literature are inappropriate. For example, Fama (1998) suggests that bad-model problems are responsible for many rejections of the EMH. He argued that the models employed by researchers do not completely explain expected returns; hence, abnormal returns are detected when such models are utilised. Malkiel (2003) presents critical reviews of all the major theories that refute the EMH. His evidence proves that most studies that criticise the EMH are unable to construct a portfolio that systematically provides risk-adjusted abnormal returns. In essence, Fama (1998) and Malkiel (2003) say that, in the event of the test of the EMH being conducted correctly, it is highly unlikely that the EMH will be refuted.[4]

Much of modern portfolio theory and economic modelling is based on assumptions of rational expectations and efficient markets. The repercussions of flawed underlying theory can have enormous adverse effects on the wealth of households, individuals and governments. In particular, the incorrect use of the underlying theory can have far-reaching consequences for various people. For example, in 2004 a tax accountant

[4] Conducted *correctly,* means that measures to account for risk and returns have been carried out correctly. Sensitivity to a number of things, such as transaction costs, dividends, market microstructure etc., is essential for the outcome of the tests. This is explored in detail below.

was sentenced to a 24-year jail term on fraud charges. The court applied the EMH to prove his guilt. On 7 September 2006, *The Economist* reported that the judge, who was responsible for the sentencing, was mistaken in his economics and the 24-year jail term was thrown out. It is essential that the EMH-theory be applied in conjunction with appropriate models, data and econometrics, which are becoming increasingly straightforward because of the increased understanding of efficient markets.

2.1.2 Behavioural Finance and Financial Anomalies

In stark contrast to the EMH, behavioural finance provides evidence of financial anomalies and systematic patterns in asset prices that can be exploited by informed investors. Such patterns are not possible if markets are efficient. Financial anomalies often imply that returns are predictable. Event studies, a sub-field within the behavioural literature, often suggest that returns are larger at certain times than at others. Two well-known anomalies are the "Holiday effect", which suggests that above average returns and below average variances occur before a holiday (Brockman and Michayluk, 1998); and the "Day-of-the-week effect", which suggests that the average returns and variances differ across the five weekdays (Jaffe and Westerfield, 1985).[5] Brav and Heaton (2000) present a comprehensive overview of competing financial anomalies. Their study shows that the two most frequent explanations of financial anomalies are investor irrationalities, which often build on cognitive biases; and structural uncertainties, which are present when investors do not know the 'true' structure of the economy (Brav and Heaton, 2000). Cognitive biases are motivated by investor psychology. Therefore, the behavioural literature that focuses on investor

[5] Average returns have been found to be lower on Mondays relative to Fridays (Jaffe and Westerfield, 1985).

irrationalities is heavily influenced by psychological and to a lesser degree sociological theories.

The increased importance of investor psychology in the finance literature is mostly due to two prominent psychologists, Amos Tversky and Daniel Kahneman. The two are immensely important in that they challenged the validity of the prevailing expected utility theory introduced by von Neumann and Morgenstern (1944). Tversky and Kahneman (1979) introduced the prospect theory, which, in its simplest form, shows that investors perceive losses to have a larger impact than gains. In finance, this translates to investors being quick to sell when their investments yield profits, but are reluctant to sell when their investments yield losses in the hope the losses turn around and become profitable (Odean, 1998).

Stock price clustering is the focal point of this thesis. The behavioural literature provides several explanations of the phenomenon of stock price clustering. A recent paper by Ikenberry and Weston (2003) proposes that a plausible explanation of stock price clustering on nickels and dimes is because of psychological biases among investors.[6] One hypothesis that is frequently employed as a leading explanation of investors' tendency to cluster on round numbers is the negotiation/price resolution hypothesis (Ball, 1985). This hypothesis is that clustering occurs as a combination of market makers who engage in a practice of rounding quotes to near-by numbers and the convenience of clustering for trimming down negotiation costs (Chung and Chiang, 2006). Kahneman, Slovic and Tversky (1982) show that the limited ability of investors to process large datasets causes a clustering towards numbers that are 'easy'

[6] In the U.S coinage system, a 5-cent coin is referred to as a "Nickel" and a 10-cent coin is referred to as a "Dime". Therefore, clustering on Nickels and Dimes means clustering on 5- and 10-cents [0, 5, 10, 15, …., 95].

to comprehend. ‘Easy’ numbers are round numbers that investors can quickly manipulate in their mind to make rapid investment decisions. A comprehensive discussion of clustering on round numbers is provided below.

Briefly mentioned above are the holiday- and the day-of-the-week effects. Studies of calendar events are referred to as event studies, which are conducted to evaluate the impact of financial ‘events’ on a firm’s stock price. MacKinlay (1997) presents a well-organised survey of methods used in financial event studies. He discusses pros and cons of most models, including the Capital Assets Pricing Model (CAPM) and the Arbitrage Pricing Theory (APT), and presents results from prominent events studies. Event studies often find evidence of abnormal returns in relation to a financial event. Such findings are often in conflict with the EMH. It should be noted that event studies are not confined to the study of seasonal effects: events, such as dividend announcements, mergers and acquisitions, stock splits or earnings announcements are also researched as event studies. It is further important to stress that this thesis is not an event study in the traditional sense in which returns are the dependent variable. Traditional event studies examine the effect of an event on asset returns and are often motivated by the belief that there are anomalous returns in relation the event. This thesis studies an event, expiration Fridays, but studies prices in levels rather than in first differences. Hence, it is an event study, but in a somewhat unusual manner.

The behavioural finance literature continues to grow and researchers are learning more about the behaviour of investors. The behavioural literature points out the shortcomings of the EMH and suggests alternative solutions to the same problems. Because of criticism from the behavioural literature, the EMH is gaining strength as weaknesses of the theory are identified and solutions are developed. Many of the

anomalies identified in the behavioural literature are simply due to flawed data (Beechey, Gruen and Vickery, 2000) or poor modelling (Fama, 1998). Although the two schools of thought are still very much divided, there are features of behavioural finance that can assist the understanding of EMH-shortcomings. Furthermore, the consumer of either theory needs to evaluate the assumptions, data, models and econometrics used in studies in order to rely on the conclusions of the research. If unexpected results are identified, then it is quite likely that errors have been made in one or more steps of the analysis.

2.1.3 Market Microstructure

Market microstructure is emerging as one of the *it*-topics to study in finance. An increasing body of literature addresses the underlying microstructure of financial markets. Market microstructure-researchers study topics, such as who participates in particular markets; the rules that investors, traders, and brokers comply with in order to participate in the market; and the period in which trading is allowed.

Two landmark studies of market microstructure are O'Hara's (1995) book, which provides a comprehensive overview of theoretical topics within microstructure research, and Harris's (2003) book that takes a more practical approach to the underlying microstructures of well-known markets. Both books take the reader on a journey through rules and legislation of markets with different microstructures. In particular, both books emphasise that students of financial markets should be aware of differences in the microstructure of market-maker markets, such as the Nasdaq, and specialist markets, such as NYSE and AMEX. Insensitivity to the microstructure of financial assets is a cause of error in empirical research. This is investigated in detail

below where a chapter is dedicated to the significance of microstructure to empirical studies.

The market-structure topics in this thesis focus on differences between Optionable Stocks (OS) and Non-Optionable Stocks (NOS); and Over-the-Counter (OTC) stocks and Listed stocks. The microstructure literature of OTC and listed stocks is large. However, the literature that studies microstructure issues between OS and NOS is far narrower, but the overlap between the two fields of study is significant. A sound understanding of institutional differences between OTC and listed stocks does enhance the understanding of differences between OS and NOS. For example, Ni, Pearson and Poteshman (2005) suggest that the distribution of OS significantly differs from that of NOS on expiration days. They conclude that this difference is due to price manipulation of the underlying assets induced by firm proprietary traders with net-written option positions who try to bring their short options to expire Out-of-the-Money (OTM). This manipulation causes the price of the underlying assets to cluster on option strike prices. Grossman, Miller, Cone, Fischel and Ross (1997) produce evidence that price-continuity- and absolute-time priority rules, which are characteristics of the NYSE/AMEX, are leading factors in lower clustering on round numbers. They also suggest that the NYSE specialist, who is responsible for the public limit-order book, is the most important structural factor that leads to lower clustering among listed stocks.[7] Such rules do not exist on Nasdaq. It is, therefore, necessary to include a control variable that separates Nasdaq from listed stocks when studying clustering among OS and NOS. This is not done in Ni et al.'s (2005) study.

[7] Grossman et al. (1997) document clustering on round numbers, in their case even-eights, which often correspond to options strike prices. They further find that this clustering is lower for NYSE stocks relative to OTC stocks. Therefore, it can be implied from their results that clustering on options strike prices is lower for listed stocks relative to OTC stocks.

Panayides (2004) and Glosten (1989) provide further discussions of the role of the NYSE specialist and the impact of the price-continuity rule on quoted prices and liquidity-creation in response to adverse selection problems.

Madhavan (2000) surveys the microstructure literature and underlines the importance of incorporating differences in microstructure between assets that are included in financial research. Whether the research is concerned with asset pricing or corporate finance, ignorance of underlying microstructure can lead to the discovery of non-existing anomalies and false rejections of the EMH. Therefore, emphasis on microstructure plays a central role in this thesis. In particular, the analysis below accounts for automatic exercise rules; whether stocks are listed or OTC; the number of market makers following the particular stock; each stock's strike-price increment; and the nature of the underlying asset.

2.1.4 Financial Derivatives

The existence of exchange-traded option contracts can be traced back to the Netherlands in the 17th century (Miller, 1995). However, it was not until 1973, when the Chicago Board Options Exchange (CBOE) decided to list 16 call options, that these securities exploded in popularity. The introduction of exchange-traded options coincided with the seminal paper by Black and Scholes (1973) in which they derived an option-pricing model that changed the path of finance. Both applications of derivative products, such as options and futures, and the academic literature have prospered since the early 1970s. Among the more influential theorists following Black and Scholes (1973) are Merton (1973), Cox, Ross and Rubinstein (1979) and Hull (1989). Merton (1973) derived an option-pricing model similar to that of Black and Scholes, which is know as the Black-Scholes-Merton model; Cox, Ross and

Rubinstein (1979) developed a simplified method to price options by employing binomial trees; and Hull (1989), surveyed the options theory and translated an otherwise mathematically rigorous literature into a language that students of financial markets can understand.

The literature is immense, whether the topic is option pricing; the use of options to reduce or increase risk; the effects of options on the underlying assets; or the microstructure of options markets. The current thesis highlights the impact of options and futures on the distribution of the underlying assets. Avellaneda and Lipkin (2003); Schwartz, Van Ness and Van Ness (2004); Ni et al. (2005); Jarrow (1994); Chung and Chiang (2006); and Krishnan and Nelken (2001) are examples of authors finding evidence that derivatives cause inefficiencies in the market for the underlying assets. On the contrary, Easley, O'Hara and Srinivas (1998); Ross (1976); Figlewski and Webb (1993); and Kumar, Sarin and Shastri (1998) produce evidence that options are efficiency-enhancing for the market of the underlying assets. Hence, it boils down to the individual student's interpretation of the above articles to decide whether derivatives are "good" or "bad" for the market of the underlying assets. It is important, however, that the methodologies employed in these studies are understood and scrutinised before any policy changes are implemented.

2.1.4.1 Expiration-day effects

A number of researchers have studied expiration-day effects. In particular, *triple witching*-days in which stock options, stock index options and stock index futures expire on the same day are widely studied.[8] The expiration-day effects that are most

[8] Triple witching happens four times a year, on the third Friday of March, June, September and December and it is an event when stock index futures, stock index options, and stock options expire on the same day.

commonly studied are price-changes (volatility) and volume-changes. Stoll and Whaley (1991) find that the economic significance of price- and volume changes on expiration days is insignificant. Their study focuses on expiration-day effects following a change of the settlement procedures of S&P 500 and NYSE index futures and index options. In 1987, the expiration of index futures and index options changed from expiring at the closing of the market, on expiration days, to expiring at the opening of the market, at expiration days. Herbst and Maberly (1990) present evidence that these changes to settlement procedures have no impact on volatility for expiration days. However, they found that some of the volatility moved intra-day, from the close to the open.

This thesis, however, studies a third dimension of expiration-day effects, namely the distribution of the underlying equity prices. Hence, this study focuses on the price-level on expiration days rather than the change in price on expiration days. Krishnan and Nelken (2001) present empirical evidence that Microsoft's closing price clusters within 25 cents of an option's strike price on expiration days relative to non-expiration days. Furthermore, they conclude that low volatility stocks are more likely to cluster. Avellaneda and Lipkin (2003) derive a theoretical model that accounts for the clustering of OS at expiration. They argue that high *open interest*, which illustrates the number of contracts that are outstanding, is a main cause of clustering.[9] In particular, stock options that, on average, have low open interest, but because of hedge funds' "impulsive" buying of such options, which causes the open interest to reach new highs, have a larger probability of clustering. Ni et al. (2005) present 'striking' evidence that OS cluster much more on expiration days compared to NOS.

[9] Cause-and-effects are difficult to establish. Therefore, high open interest is potentially a symptom of clustering and not a cause.

They provide two explanations for the clustering. First, investors who have net purchased option-positions push the underlying asset-price towards the clustering strike price by rebalancing their delta-hedge. Secondly, investors, mainly market makers and firm proprietary traders, manipulate the underlying asset price in order to manoeuvre their net-written option positions from expiring ITM. There is a comprehensive discussion of Ni et al's (2005) analysis below. Furthermore, their analysis is replicated, but this analysis is augmented with greater sensitivity to the underlying market microstructure. The results of this analysis are presented in tables below (Table 3, Table 4, Table 5 and Table 6).

Chapter 3: Methodology

3.0 Data collection and the integrity of social science data

The data, tests and definitions in this section are relatively broad and explanations of more specific terminology are provided in the context in which they are required.

The major limitation of this research was to gain access to appropriate data. Although official closing prices are readily available at several data vendors, options data are harder to access. Five data vendors are used to collect the relevant data. These are Yahoo finance, Standard and Poor's, NASDAQ, MarketWatch and Datastream.[10] Historical options data are not publicly available and the purchase of such data is beyond the budget of this research. Options data are only accessible to the public a few hours between the market closes on every trading-day, until it re-opens on the following day. Because the particular focus of this thesis is on expiration-day effects, the data were collected in the hours after the market closed on expiration-days. Each option-chain was downloaded one-by-one and keyed into a spreadsheet by hand. The data were collected for the months of February through August 2006.

The data contain information about the contract-symbols; the open interest for put- and call contracts; the volume for each contract; the bid-ask spreads; the last trade price; and the change in the last trade price. The data also contain a price that, in the current thesis, is referred to as the "4.00 pm EST.-price".[11] Ambiguity surrounds the actual meaning of this price and it is unclear what it represents. This is discussed further below. The options data were downloaded from Yahoo Finance and

[10] Datastream is not a publicly available database. However, data from Datastream are only used as a robustness check of the publicly available data and are not the primary source of data. MarketWatch is operated by Dow Jones.

[11] It is important to note that this price represents Eastern Standard Time (EST.), which is the New York-time.

MarketWatch. Because of time-constraints, the sample was restricted to stocks in the S&P 100, NASDAQ 100 and S&P 600 indices. This ensured that both mega-cap and low-cap stocks were included, which enabled the investigation of differences between "large" stocks and "small" stocks.

The S&P 600 is particularly interesting because, roughly, half of the stocks are optionable and half are not. This facilitates a comparison of stock price clustering for optionable and non-optionable stocks with similar characteristics. Ikenberry and Weston (2003) show that clustering is inversely related to volume and stock price.[12] For this reason, the inclusion of S&P 600-stocks alongside the high-cap stocks biases the results to favour clustering for the whole sample.[13] Furthermore, the fact that none of the mega-cap stocks is non-optionable makes it extremely difficult to test the effect of options on expiration-days for these stocks.

Above, the 4.00 pm EST.-price was introduced. This price is the price that is posted on most data vendors' websites in the hours after the market closes on expiration days. It is implied that this represents the closing price.[14] The closing price on expiration days is particularly important to investors who trade options, because this price determines an option's *moneyness*, which is a measure of cash flows (profits) that the option generates, at expiration (Hull, 2003).

[12] Given that Ikenberry et al.'s (2003) results are not subject to the DeLong and Lang (1992) hypothesis (described below), the results imply that clustering on strike prices should be lower for high-cap, high-volume stocks.
[13] The volume of S&P 600 stocks is much lower than that of mega-cap stocks.
[14] There is ambiguity surrounding the official closing price. In particular, the fact that several data vendors provide different closing prices reveals that there are data inconsistencies that cause confusion among market participants. It is important to note that this research employs Standard and Poor's closing price as the official closing price. This is because, with the benefit of hindsight, Standard and Poor's provide closing prices that correspond to those used by the OCC in determining moneyness.

"Normal" investors, who are not institutional investors or market makers, have cut-off times, in which they need to decide whether they are going to exercise their options, around 4.30 pm EST[15] At this time, the only price that these investors have access to is the 4.00 pm EST-price. However, there is a major problem with this particular price: It is not the official closing price. The official closing price is the price that the moneyness is determined from, but this price is not in the public domain until hours after the cut-off.

During the current research, many inconsistencies between the 4.00 pm EST-price and the official closing price were identified. Catania and Maberly (2006:2) identify incidents in which the 4.00 pm EST-price is inside the automatic exercise thresholds (~ATM), but the official closing price is ITM. Such incidents can create problems for investors: If for example, an investor who is long the call, but does not have the cash to be exercised, believes that his option is not going to be exercised, then he is not going to try to sell his option. However, after his cut-off time, he discovers that his option is ITM. At this point in time, there is no market for him to sell his call and he is subject to a margin call that he cannot meet. Therefore, the discrepancy between the price that investors can see and the price that really matters for expiring options is material.

Furthermore, Catania and Maberly (2006:1) provide evidence of inconsistencies in reported closing prices between different data vendors. In some cases, the inconsistencies have economic impact on conclusions drawn by researchers. The

[15] Charles Schwab and Ameritrade, America's largest and third largest online brokers, respectively, require customers to submit their exercise instructions before 4.30 EST on expiration Fridays (Keoun, 2006). The following information was retrieved from Ameritrade's website:
'If you do not want your long option exercised, contact us by 4:30 pm EST on the last trading day for the options contracts'. – Ameritrade 7 November 2006

reliability of data is always an issue when conducting social science research and this is the reason why several vendors are employed in this thesis. For every inconsistency that was identified, the other vendors were checked and the correct official closing price was included. Hence, there is a chance that this research is based on flawed data, but by using five different data-sources, the risk is minimised. Furthermore, it appears that the closing prices that are provided by Standard and Poor's correspond to the official closing price that the Options Clearing Corporation (OCC) uses to determine options' moneyness. However, these prices are not available until several hours after the cut-off times for all investors.

Although data are of concern throughout this research, another and perhaps more significant problem is present. That is, the credibility of existing economic literature and the power of the hypotheses tested. De Long and Lang (1992) question the reliability of economic null hypotheses. They argue that economists choose nulls that are easy to reject and, consequently, the confidence in the null is reduced if it is rejected; and that standard statistical methods do not allow researchers to formulate statements about the sample distribution and whether the null hypotheses are true or false. The users of economic literature, therefore, have to be critical of rejected null hypotheses. Conclusions based on statistical testing should be carefully scrutinised before they are accepted as facts. In order to maximise the credibility of the hypotheses tested, this thesis employs conservative null hypotheses that, if anything, bias the rejection decision to favour the results that are under scrutiny.[16]

[16] An explanation of this point is provided below. However, this thesis hypothesises that optionable stocks are more efficient than non-optionable stocks in the sense that they follow the uniform distribution more closely relative to non-optionable stocks. The uniform distribution is the "conservative" null hypothesis and this is explained in detail below.

3.1 Tests and Robustness

Two major hypotheses are tested in this thesis. First, a test of whether the distribution of stock prices is uniform across given intervals. These intervals can be decimals, which are the smallest unit in which stocks can be exchanged, or larger intervals that carry more economic information. Secondly, a test to identify whether there is a larger frequency of closing prices in the area near an option strike price relative to the uniform distribution is performed. In particular, frequencies of clustering within 15- and 25-cents of option strike prices are compared to the expected, uniform, frequencies on these intervals.

Because a single test statistic may reject the null hypothesis when, in fact, it is true (type I error), at least two tests were employed to minimise the risk of type I errors. To test the goodness-of-fit between the observed data and the uniform distribution, both Pearson's Chi-square test and a log-ratio Chi-square-test, which is based on relative entropy, were employed. Pearson's Chi-square is given by:

$$P = \sum \frac{(O_i - E_i)^2}{E_i} \sim \chi^2_{n-1},$$

where O_i is the observed clustering on interval i and E_i is the expected clustering on interval i. The log-ratio Chi-square test statistic is given by:

$$LR = 2n \sum p_i \ln(p_i / q_i) \sim \chi^2_{n-1},$$

which is Chi-square distributed at n-1 degrees of freedom. p_i is the observed frequency on interval i and q_i is the expected frequency on interval i. A discussion on the usage and derivation of this test is can be found in Jessop (1995). These tests were

employed to examine whether the observed clustering follows the hypothesised uniform distribution.

On occasions, the significant levels/p-values, which rejection decisions are based on, differ from one test to another. This is particularly prominent when the different tests employ different distributional assumptions. However, in this research, there were no major problems of this kind. Although the significance levels varied slightly, decisions to reject the null hypothesis at the 5%-level was relatively consistent across tests.

To test for clustering around options strike prices, the expected clustering was calculated as a function of options' strike deltas. This calculation is described in detail below. The observed clustering was then tested against the expected clustering. Two test statistics were employed. First, the large-sample test of hypothesis about a population proportion was used. This is a Z-test and is given by:

$$Z = \frac{\hat{p} - p_0}{\sigma_{\hat{p}}},$$

where,

$$\sigma_{\hat{p}} = \sqrt{\frac{p_0 q_0}{n}},$$

and,

$$\hat{p} = \frac{\sum_{i=1}^{n} C_i}{n} \text{ and } q_0 = 1 - p_0,$$

where,

$$C_i = \begin{matrix} 1 \; if \; S_i^{FRI} \leq |K_i \pm clustering \; interval| \\ 0 \; otherwise \end{matrix}.$$

Where the symbols represent:

$\hat{p}$ = the observed clustering

p_0 = the hypothesised/expected clustering

$\sigma_{\hat{p}}$ = the sample standard deviation

$q_0 = (1 - p_0)$

C_i = zero-one dummy that equals one if the stock price is within a given interval from the option strike price.

Secondly, the group mean comparison test, which is a t-test, was employed.[17] This test is comparable to the Z-test, but it assumes the t-distribution as opposed to the normal distribution, which is assumed by the Z-test. The t-distribution has "fatter" tails than the normal distribution and rejection decisions can vary when the number of observations is low.

[17] The t-test is built into STATA, which is the primary statistical software utilised in this thesis. STATA-codes for all statistical analyses carried out in this thesis are provided in Appendix 6.

Chapter 4.0: Results

Question 1: Are there market microstructure differences between OTC, AMEX and NYSE stocks; if there are, how do they influence research of stock prices behaviour on option expiration days?

"*Market structure determines what traders can do and what they can know*"
- Larry Harris (2003, p. 7)

In order to perform a coherent analysis of the behaviour of stock prices on expiration days, it is crucial to understand the microstructure of the relevant market. The microstructure of a particular market consists of a number of things, such as rules and regulations; information-flows; market participants; and institutional details. In this paper, the key area of interest is the different stock price behaviour between optionable stocks and non-optionable stocks with similar attributes, such as market capitalisation, volume and volatility[18]. Therefore, it is important to identify differences in microstructure between the markets for optionable stocks as opposed to non-optionable stocks. Furthermore, attention has to be directed to institutional differences between listed stocks, as opposed to over the counter (OTC) stocks, and the ways they are traded. A large proportion of firms in the dataset are OTC, which are stocks that trade on the Nasdaq. Knowledge about automatic exercise rules; strike-price increments; closing-price accuracy; and the composition of the underlying assets (common stocks or Exchange-Traded Funds (ETF)) is fundamental to a sound analysis of stock price behaviour. Unfortunately, there is limited published work on options-market microstructure and some of the rules that constitute options-markets are "hidden" away in regulators' documents. Knowledge of such rules can be

[18] For the clustering analysis of optionable stocks, the main source of data is the S&P 600. This index caters for all the criteria stated desirable. Particularly suitable is the fact that approximately half of the stocks in the 600 are optionable and half are not. Furthermore, the stocks in this index have similar characteristics, which accommodates for a comparative analysis.

acquired after years of trading experience. Lack of knowledge about the relevant microstructure can lead to biased and spurious analyses.

4.1 Automatic exercise

Automatic exercise is a mechanism that protects the option holder by making sure that options that are in-the-money (ITM), by a certain threshold, are exercised. The Options Clearing Corporation (OCC) has set three thresholds for automatic exercise.[19] These are:

1. Equities: 15 cents ITM options for *firm and market maker accounts*
2. Equities: 25 cents ITM options for *other investors (customers)*
3. Index options: 1 cent ITM for all accounts.

It is important to note that these thresholds have changed over time. The above thresholds changed in September 2004. Before that, the threshold for customer accounts was 75 cents ITM and for market maker and firm accounts, the threshold was 25 cents ITM.[20] It should be noted that as of 20 October 2006, the automatic exercise thresholds change from 25 cents and 15 cents for the customer, and firm and market maker accounts, respectively, to 5 cents for all accounts. The effect of these changes can only be studied after a sufficient period with new thresholds has passed. However, the two intervals, [K ± 14] and [K ± 24], constitute the thresholds in which no automatic exercise take place for the sample employed in this thesis.[21] Table 1

[19] This information was gathered from the Chicago Board Options Exchange's (CBOE) website: http://www.cboe.com/LearnCenter/Concepts/Basics/terminology.aspx

[20] It is important to be aware of differences between customers and market makers, as the two groups are not on an entirely level plying field; for example, market makers have different cut-off times, which extend beyond the deadline for ordinary customers, to make exercise decisions.

[21] Note that we are excluding index options, which implies that the two remaining intervals are [k ± 14bps] and [k ± 24bps].

(below) illustrates a few scenarios in which the automatic exercise differs for customer and market maker accounts, respectively.

Table 1: Summarises three examples of automatic exercise given different levels of moneyness. A call option is In-The-Money (ITM) if $S_0 > K$ and $S_0 < K$ for a put option. Hence, ITM options lead to a positive money flow if the options are exercised.[22] Conversely, Out of The Money (OTM) options are generally not exercised and the owner of the option loses the premium paid for the option.

	A	B	C
Price	$ 19.78	$ 25.25	$ 25.24
Strike	$ 20.00	$ 25.00	$ 25.00
Put	$ 0.22 ITM	OTM	OTM
Call	OTM	$ 0.25 ITM	$ 0.24 ITM
Market maker automatic exercise	YES	YES	YES
Customer automatic exercise	NO	YES	NO

In example A, the spot price is $19.78, which is 22 cents below the $20 clustering strike price. Consequently, the $20 put is 22 cents ITM, but is not subject to automatic exercise in the customers' account. Because the option is not automatically exercised, the probability of the option being exercised is reduced. Ni et al. (2005) suggest that investors with net written options positions (investors that sell options) manipulate the underlying stock price to the point at which they expire OTM. They do not discuss the effects of automatic exercise. This thesis, however, argues that sensitivity to automatic exercise is important and that the incentives to manipulate are non-linear the closer to ATM an option becomes. The proportion of options that are exercised increases the deeper ITM an option goes. Alternatively, options that are ITM by only a small portion are less likely to be exercised. Therefore, once the option is less then 25-cents ITM, the incentives to manipulate the underlying stock are fading.

[22] Transaction costs require an option to be ITM by as much or more than the transaction costs before the option becomes profitable.

In example B, an investor buys a call option with a strike price, K, of $25. Further, the closing price on expiration Friday is $25.25. In this case, the option is 25 cents ITM and automatic exercise is implemented. If the closing price is $25.24, as in example C, the option is not subject to automatic exercise in the customer accounts. When an option is ITM, but not qualified for automatic exercise, two things can happen: either the option is left unexercised, or the option purchasers can instruct their broker to exercise the option. In the case in which an ITM option is not exercised, the option-writer gains because the losses that are owed are not claimed.

Automatic exercise is important, as it is likely that ITM options that are not automatically exercised are not exercised at all. After personal correspondence with the OCC, I have been informed that 70% of options that are $0.05 - $0.24 ITM are exercised.[23] Ni et al. (2005) imply that that this figure is kept artificially low by market manipulation. To my knowledge, there is no publicly available information about the proportion of unexercised ITM options at every decimal between $0.05 and $0.24. Naturally, a larger proportion of options are expected to be exercised, the deeper ITM the option becomes; which is what this research will assume for the remaining discussion.

[23] The 70% proportion was determined from data collected on a 18 month interval (June 2004 through December 2005). This information can be obtained from the SEC (http://www.sec.gov/rules/sro/occ/2006/34-54306.pdf) and the OCC (http://www.optionsclearing.com/publications/rules/proposed_changes/sr_occ_06_05.pdf).

4.2 Strike price increments

The key assumption of this thesis is that prices are uniformly distributed across intervals of 15 and 25 basis points. Tests of this assumption are performed below. Because of the uniform distribution's desirable properties, such as a horizontal density function, it becomes straightforward to calculate the expected clustering.[24] It is instrumental to calculate an expected clustering, as an efficiency-benchmark, to determine whether clustering on expiration days is efficiency enhancing or destructive.[25]

In order to calculate the expected clustering, it is necessary to identify each firm's strike delta, ΔK_i. Figure 1, below, illustrates a case where a company's ΔK_i = \$2.50 (250 bps). The expected clustering, for a given interval, is simply the ratio of the interval that defines the clustering divided by the strike delta. This implies that for a clustering interval [$K_i - 24, K_i + 24$], the expected clustering is:

$$E(C_i) = \frac{49}{250} = 19.6\% .$$

Hence, given that the stock price is on an interval, for example [21.26, 23.75], the probability of the stock price being on the $K_i - 24, K_i + 24$ -interval (shaded area) at the close is 19.6%. Had ΔK_i been different from \$2.50, the expected clustering would have been different, which is why it is critical to identify each firm's strike delta. Krishnan and Nelken (2001), Avellaneda and Lipkin (2003), and Ni et al. (2005)

[24] The horizontal density function of the uniform distribution accommodates for easy calculations of the expected clustering.

[25] It is known from Ikenberry et al. (2003) that the distribution is not uniform across every decimal, but the analysis below finds that the distribution is increasingly uniform at larger and more economically meaningful intervals. The subsequent chapter discusses the uniformity assumption in detail.

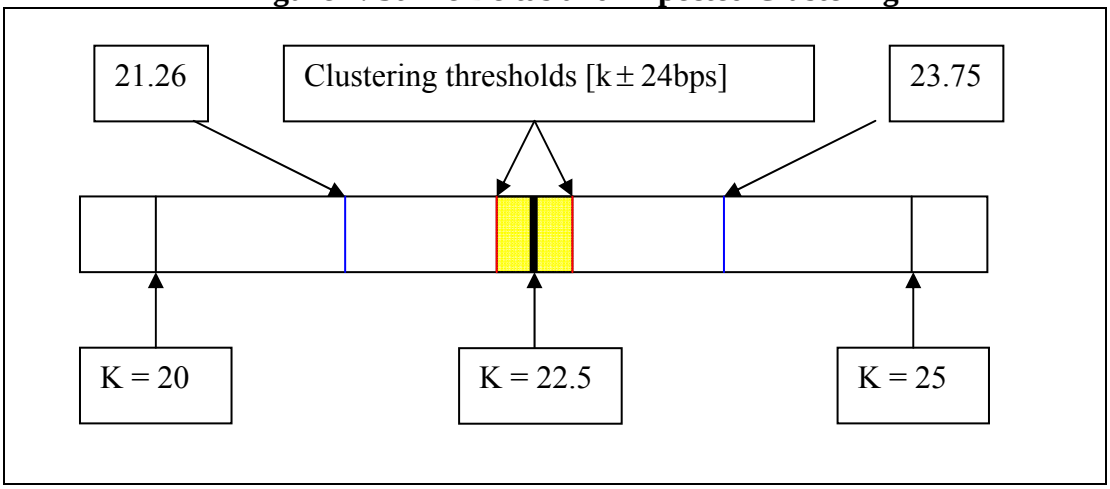
demonstrated that there is a move towards option strike prices on expiration Fridays.[26] Commonly, these authors lean toward the conclusion that the altered distribution is caused by inefficiencies in the market. Most noteworthy is Ni et al's (2005) article that concludes that the "push" towards the strike price is being caused by market makers and proprietary traders who manipulate underlying asset prices.

Neither of the articles described above employ any theory to establish the expected clustering, yet, they manage to conclude that the expiration day clustering is a "bad" event. This research proposes a way to calculate the expected clustering based on the clustering thresholds and individual stocks' strike deltas to answer the question: On expiration days, is the observed clustering closer to its expected value (that is more efficient) relative to non-expiration days? If it is closer to the expected value, this can be attributable to the enhanced trading volume and increased number of products changing hands on expiration days. An additional exploration of this avenue is performed by studying triple witching days in separation from non-witching days. Triple witching happens four times a year, on the third Friday of March, June, September and December and it is an event when stock index futures, stock index options, and stock options expire on the same day. Triple witching is, because of the increased trading-activity, the expiration in which the observed clustering should be the closest match to the expected clustering. The data for this thesis cover two triple witching Fridays, March and June, which are points of reference as far as efficiency is concerned; because witching days experience much larger trading volumes than non-witching days and, with the exception of the opening on witching days, theory

[26] Krishnan and Nelken (2001) explore how Microsoft tends to *cluster* on expiration days relative to non-expiration days. Avellaneda and Lipkin (2003) perform an analytical analysis of clustering where they conclude that, given 'certain circumstances', stock price clustering can happen due to market makers' delta hedging.

suggests that increased trading activity increases information flows.[27] Market participants can make better decisions given the increased information and, consequently, efficiency should be improved.

Figure 1: **Strike Deltas and Expected Clustering**



This figure illustrates a "typical" scenario, where the clustering strike price is $22.50 and the strike delta is $2.50. In this case, the red lines mark the outer boundaries for the K ± 24bps clustering interval. The blue lines constitute the interval in which K=$22.50 is the clustering strike price. For example, if the share price is less than $21.26, then the clustering strike price becomes $20. Hence, given that the share price is on the [21.26, 23.75] interval, the expected clustering is 19.6%.

[27] S&P 500 futures and options are settled based on the *opening* price of the underlying asset (Barclay, Hendershott and Jones, 2006). Because arbitrageurs unwind (buy and sell) their index positions on the expiration-day opening, liquidity shocks cause extreme increases in volume with no additional information to market participants (Stoll and Whaley, 1990).

4.3 The underlying assets

Ni et al. (2005) claim that roughly 2500 stocks are optionable, which is true on average. However, they provide little information about the usage or break-up of their data. It is inappropriate to include all underlying assets that have stock options trading on them. This is the case because the underlying asset of an equity option consists of several assets types; these are common stocks, Exchange Traded Funds (ETF), trusts, American Depository Receipts (ADR) and holding companies. If the underlying asset prices are being manipulated, it is important to distinguish between the types of underlying assets.

Most significant is the case of ETFs. It would be unattainable for any market maker, or a pool of market makers, to manipulate the underlying asset price if the underlying asset is an ETF. An ETF is an index fund that trades on exchanges like a stock (Fuhr, 2001). The strong laws of arbitrage govern trading in these funds. This complicates the job of manipulating the price, because the manipulator would have to manipulate all the shares in the underlying index. One example is the NASDAQ 100-trust shares (QQQQ).[28] There are 100 high-cap stocks included in the QQQQ and they would all need to be manipulated to force the options OTM, which is virtually an impossible task to perform without being detected and penalised. Furthermore, the costs of such a scheme would be immense and, presumably, not worth the risk. It is difficult to imagine any underlying asset, apart from common stocks and ADRs that, in theory, can be subject to manipulation. Therefore, it should be noted that only about 2200 common stocks and 200 ADRs are optionable.[29]

[28] QQQQ is the stock ticker symbol of the Nasdaq 100-trust shares.

[29] According to Standard and Poor's, an ADR is "a certificate representing home market securities and provides U.S. investors with a convenient way to invest in non-U.S. securities without leaving the

The analyses carried out in this research include common stock and ADRs only. If ETFs are included in the analysis, the estimates will be biased in the sense that they may be considered to be manipulated, when, in fact, no manipulation can happen given their unique qualities. It is conceivable that strong players in the market can manipulate individual stocks. It is, however, less plausible that such manipulation is allowed to happen as frequently and systematically as suggested by Ni et al. (2005). There are strict rules against market manipulation and if detected, severe penalties, such as fines, suspension, expulsion from trading, or imprisonment can be imposed.

4.3.1 Stealth Options

A *Stealth Option* is another element that researchers have to be sensitive to when analysing clustering of optionable stocks. Because stealth options do not feature in previous research, it is necessary to define the term:

"A Stealth Option is defined as an Optionable Stock that has zero volume on the strike price that is closest to the closing price. Therefore, if a stock with a stealth option clusters on expiration Fridays, this clustering is referred to as *Stealth Clustering*".[30]

If options are the explanatory factor of clustering, then there must be trading activity causing the push/drag towards the strike price. In many cases, stocks have options written on them, but the fact that there is no volume effectively makes the stock non-

---

country." Hence, it is a way of making investments into foreign companies convenient for American investors.

[30] On expiration days, 533 (13%) stocks among the optionable stocks had zero volume. Across the whole dataset, which comprises seven expiration Fridays, 1539 optionable stocks had no options volume, which was 12.95% of all optionable stocks. Furthermore, it should be noted that another definition of stealth clustering was formed; namely, that stealth clustering is the occasion in which optionable stocks have zero volume and zero open interest on expiration Fridays. This definition made no material impact on the hypothesis tests.

optionable. Following Ni et al's (2005) logic, the manipulator trades in the underlying asset in order to prevent the options from being exercised. The absence of option volume translates to zero trading, which implies that there are no incentives for the proprietary traders to manipulate the underlying. Hence, if there is no volume at the clustering strike price, then it is unwarranted to blame the clustering on optionability. Consequently, stocks that are optionable, but have no volume on the clustering strike, are treated as non-optionable stocks in the subsequent analyses.[30] However, by treating stealth-clustering stocks as non-optionable, Ni et al's (2005) results should be further emphasised. This is because the stocks with volume ought to exhibit strong clustering if optionability is the factor explaining clustering.

4.4 Summary

Numerous components make up the market microstructure for a particular market. Although a number of microstructure variables were left out in the above discussion, the factors that were perceived to be essential for the clustering-analysis were included. In particular, strike price increments (strike deltas) are important in order to establish each security's expected clustering. Furthermore, the concept of stealth clustering, which is original of this research, is defined. By ignoring any of these factors, the risk of biased analyses increases. Microstructure variables that are important for the subsequent analyses are discussed in the context in which they are required.

Question 2: What distribution constitutes the expected clustering of stock prices?

The foremost assumption of this thesis is that stock prices are evenly distributed across all given intervals. In other words, given that there are 100 bins (decimals) in a dollar, the expected frequency of stock (closing) prices in each bin is 1%. However, for the following analysis, it is not necessary that the distribution is uniform across every decimal, but rather on larger intervals.[31] In particular, the research question is: Do observed stock prices follow the uniform distribution on intervals of 15 basis points and 25 basis points? These intervals constitute the automatic exercise thresholds, which were discussed in detail above.

Since the 1930s (Gann, 1930) researchers have found a particular bias in the behaviour of stock prices, which suggests that investors tend to cluster on round numbers. After decimal trading was introduced, clustering on 5 and 10 cents has been most common.[32] Others who have documented such behaviour are Niederhoffer (1966), Harris (2003), Sonnemans (2005) and Ikenberry et al. (2003). Several reasons for clustering on round numbers have been theorised: Limit orders, in which investors can submit a buy (sell) order at a given price, at round numbers are one of the main hypotheses for clustering. Mitchell (2001) provides an excellent survey of literature suggesting why people prefer particular numbers to others. One theory of clustering is the price resolution hypothesis, which was introduced by Ball, Torous and Tschoegl (1985) and it suggests that clustering on round numbers is an increasing function of uncertainty and price volatility. Harris (1991) confirms that volatility and price

[31] The two clustering intervals, 15 and 25 cents, constitute the intervals for which the observed distributions are expected to behave uniformly. At the decimal level, the observed distributions are significantly different from the uniform, but on larger intervals, the observed distributions become close to the uniform (see section 3.1.2).

[32] In 2001, the minimum price increments (ticks) that American shares could be traded were changed to one decimal. Prior to decimal trading, price increments were most commonly in fractions of 1/8th and 1/16th.

clustering on round numbers are positively related. He further suggests that traders frequently trade at round numbers to avoid negotiation costs, which gives birth to the price/negotiation resolution hypothesis. Grossman, Miller, Cone, Fischel and Ross (1997) applaud that negotiation costs explain clustering, because negotiation for a "better" price may not be worth the costs of doing so. They extend the ongoing discussion with an analysis into the different microstructures of the NYSE, which is a specialist market, and competitive market-maker markets, such as the Nasdaq and the LSE. They find that *rule 104*, which requires the NYSE-specialist to provide continuously updated quotes (Panayides, 2004), caused less clustering than what is found in competitive market-maker markets. Consistent with their theory, they found lower clustering on the NYSE.

The behavioural literature suggests that certain cognitive biases reduce investors' ability to process large price grids and, consequently, are attracted to numbers that are "easy" to comprehend (Kahneman, Slovic and Tversky, 1982). Mitchell (2001) shows that clustering on round numbers is caused by the methods investors use to process data. Hence, the limited cognitive capacity of investors to decompose and process information leads to a preference for round numbers.

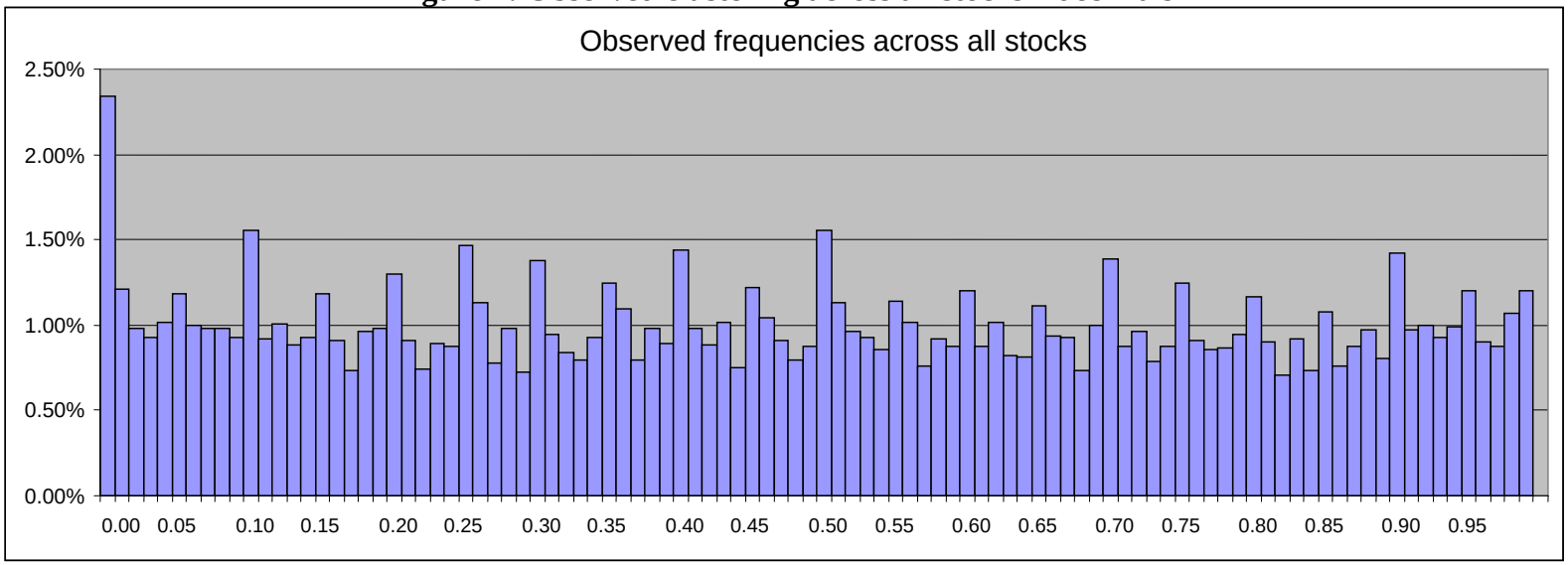
4.5 The evidence

Ikenberry et al. (2003) find clustering on nickels and dimes (i.e., 5- and 10 cents, respectively), which prevents the distribution from being uniform across decimals. I replicate Ikenberry et al.'s (2003) analysis across all stocks and report similar finding. I make no distinction between optionable and non-optionable stocks or between expiration and non-expiration Fridays. The observed distribution generated from this analysis is presented in Figure 2 below.

When studying Figure 2, it becomes obvious that there were tendencies for stock prices to cluster on round numbers. In particular, the clustering in the zero-bin was immense. The clustering in the zero-bin was so prevalent that the Chi-square value for this data point, individually, was enough to reject the uniform distribution, given 99 degrees of freedom.[33] The Chi-square test statistic for the goodness-of-fit of the observed distribution relative to the uniform distribution was 880.89 and the corresponding p-value was 0.000, which led to the rejection of the uniform distribution at the decimal level.

[33] It is not valid to test each data-point separately with the Chi-square test, but it is informative here as it indicates the strength of the clustering in the zero-bin.

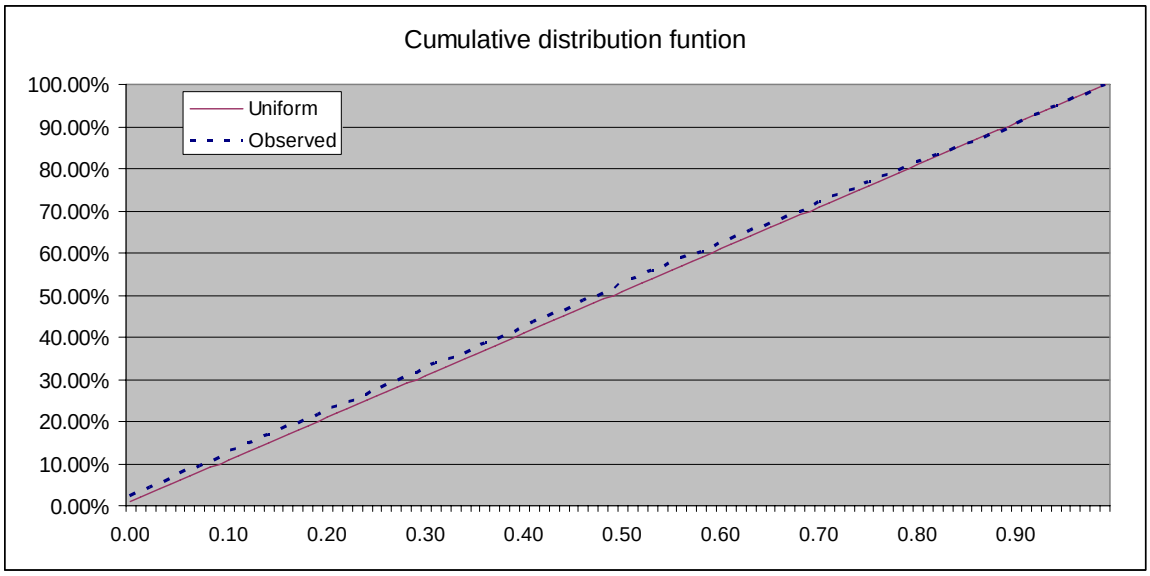
Figure 2: Observed clustering across all stocks – decimals



This figure illustrates the observed frequency distribution across decimals of every observation in the dataset, which includes S&P 100, NASDAQ 100, and S&P 600 stocks on the SECOND-, THIRD- and FOURTH Friday of every month from February 2006 including August 2006. It is clear that the zero-bin experienced a much higher degree of clustering than other bins. The three bins that experienced the highest degree of clustering were zero, 10 cents and 50 cents, respectively. The Chi-square statistic (Chi-square = 880.89) for the zero-bin alone, was large enough to reject the uniform distribution overall. When the intervals are decimals, the data are noisy and the uniform assumption is being challenged by statistical tests. The discussion below focuses on extending the analysis to include wider intervals. This will not only reduce the noise, but it makes more economic sense to match the interval-width to the automatic exercise intervals. A similar, but more detailed analysis of S&P 600 stocks is included in Appendix 4.

As discussed above, the distribution on larger intervals is more interesting because of the automatic exercise regime. The goodness-of-fit is illustrated in Figure 3, in which the cumulative distribution functions of the uniform- and observed distributions are graphed. Although the statistical difference between the two distributions is significant, the economic difference seems to be minimal.

Figure 3: The Cumulative Distribution Function



This figure illustrates the cumulative distribution functions of the uniform and observed distributions. This graph is only meant to summarise the relationship between the two distributions. It is not a test to establish statistical significance of the difference between them. However, it illustrates that the economic difference between the two distributions is minimal. The data used for this graph include all observations from February 2006 up to and including August 2006.

4.5.1 Wider intervals

Automatic exercise happens at 15 cents for market makers and 25 cents for customer accounts. The 25-cent interval is practical because it facilitates a straightforward division of the data, in which the 100 decimal-bins can be broken down to four intervals, each containing 25 decimal-bins. The expected clustering on these intervals is 25% (25 bins/100 bins) and Figure 4 summarises the observed clustering and the expected clustering. Figure 4 illustrates that the observed clustering was very close to the expected clustering. However, because of the zero-bias, the first bar, which

contains the first 25 decimal-bins, experienced stronger clustering than the residual three intervals. The Chi-square value was 17.36 and the corresponding p-value was 0.001 (Table 2, below, summarises statistical computations), which, again, led to the statistical rejection of the uniform distribution. Because the clustering in the zero- and 50-cent bins is more frequent, relative to other decimals, there is a natural clustering on zero cents and 50 cents. Consequently, because strike prices generally occur on zero cents and 50 cents, there will be a tendency to cluster on strike prices, which is what Ni et al. (2005) found.

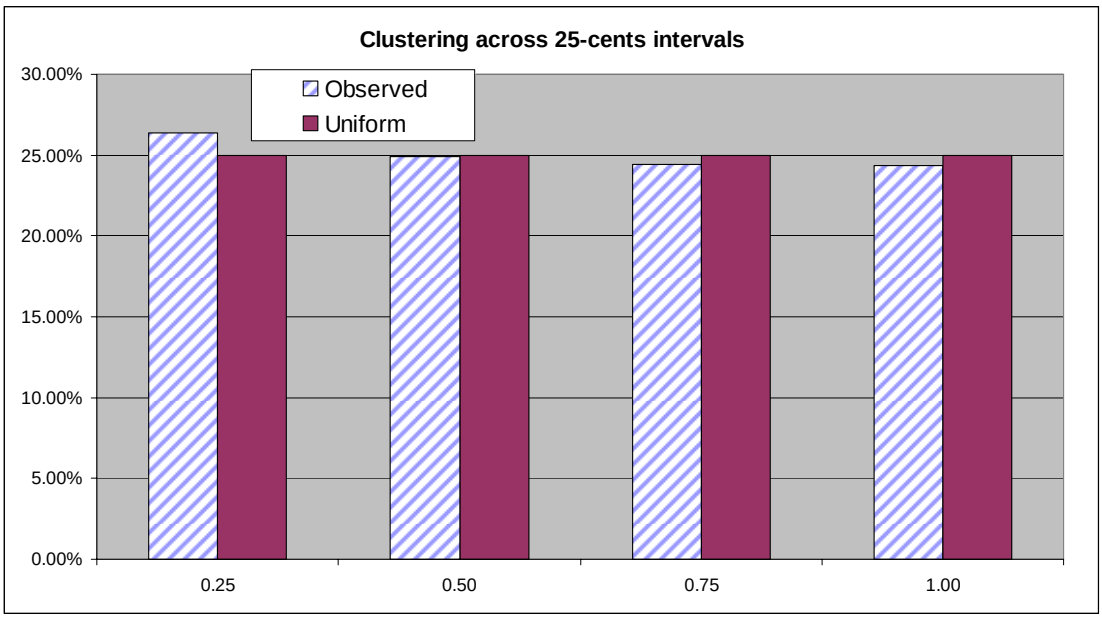


Figure 4: Illustrates the observed clustering and the expected clustering. Economically, the difference between the observed bars and the expected bars is insignificant. However, the statistical difference is significant, even at the 1% level (p-value = 0.001). The 0.25, 0.50, 0.75, and 1.00-bars contain the [0, 24], [25, 49], [50, 74], and [75, 99]-decimal-bins, respectively. Hence, there are 25 decimal points on each interval. In an options-context, the [0, 24]-interval represents the interval in which calls are marginally ITM, given that the strike price falls in the zero-bin. Conversely, the [75, 99] interval represents the ITM puts. The sample consists of the whole dataset, which includes S&P 100, NASDAQ 100, and S&P 600 stocks on the SECOND-, THIRD- and FOURTH Friday of every month from February 2006 including August 2006.

It is of interest to determine the influence of zero-bin clustering upon the rejection decision. If the uniformity assumption is rejected solely because of excess clustering in the zero-bin, the results are fragile. In particular, small reductions in the zero-bin clustering can potentially produce an increasingly uniform distribution. To test the

robustness of the above results, the data were manipulated in two different ways. First, the zero-bin observations were equally weighted across the four intervals.[34] This changed the Chi-square test statistic to 1.84 and the corresponding p-value to 0.6057. Secondly, the zero-bin observations were replaced by the expected value (166.24), which changed the Chi-square test statistic to 2.81 (p-value = 0.422) (Table 2 summarises the statistics for all relevant sub-samples).[35] These results emphasise that clustering in the zero-cent bin is the factor that independently causes the rejection of the uniform distribution across the 25-cent intervals. These results do not prove that there is no clustering, but they emphasise the lack of robustness when rejecting the null hypothesis. Because the observed data are approximately uniform, it motivates the usage of the uniform distribution as the central null hypothesis.

It should be noted that the uniform distribution is the most conservative measure of the expected distribution. Because of the natural tendency to cluster on round numbers, hypotheses tests against the uniform distribution will be biased towards rejection. This happens because the uniform distribution is the unconditionally expected distribution, but the empirics suggest that the observed clustering is larger than the uniform distribution on round numbers. As a result, for cases in which the a priori expectation is that the distribution is uniform, the intervals that include the high-frequency numbers (0, 10 and 50) are likely to exceed the uniform distribution. Therefore, if anything, the bias will support the rejection of null hypothesis. Consequently, when the expected probability of clustering is calculated, the observed frequency of clustering in the zero- and 50-cent bins is larger than the expected, which

[34] This test was performed by adding one-fourth (98.25 observations) of the 393 zero-bin observations to each of the 4 bins. It is important to note that 98.25 replaced the 393 zero-bin observations on the [0, 24]-interval.
[35] This test replaced the 393 zero-bin observations with the expected number of observations (167.68), according to the uniform distribution.

biases the test statistic of the observed distribution being equal to the expected distribution towards rejection. Because of the clustering in the zero- and 50-cent bins, there is a natural clustering on strike prices across the majority of stocks when assuming uniformity. Therefore, it is natural to expect that the narrower the intervals become, the stronger the clustering on option strike prices.[36]

4.6 Optionability and the uniform distribution

The clustering on round numbers, in the decimal system, is obvious (Figure 2). The current section investigates whether the clustering is different for Optionable Stocks (OS) relative to Non-Optionable Stocks (NOS). The primary expectation for this thesis is that OS carry more information than NOS and, therefore, should behave accordingly. That is, they should follow the theoretical clustering more closely than NOS. Ikenberry et al. (2003) did not separate OS from NOS and the contribution from the current paper is a separate analysis of the two groups of stocks. The sample was divided into four sections. First, the S&P 600 was split into two sections; one with all the OS and another with the NOS.[37] Figure 5 shows the decimal clustering for OS and NOS in the S&P 600. Secondly, the Nasdaq- and S&P 100 were analysed individually. Figure 6 illustrates the distribution for the S&P 100 and the Nasdaq 100.[38]

[36] The noise on round numbers can be reduced by increasing the intervals. However, because the zero and 10-cent bins are the two bins with the highest frequencies, any interval that includes both of these bins will have a larger clustering than other intervals. Figure 5 illustrates this in which the 25-cent bar includes both the zero- and 10-cent bins.
Furthermore, the introduction of new automatic exercise intervals (in effect from 20 October 2006), in which 5-cent ITM options will be subject to automatic exercise, should enhance strike price clustering.
[37] The thesis appendix continues the analysis of the S&P 600 in which the data are divided into four sub samples: OS before and after expiration; OS at expiration; NOS before and after expiration; and NOS at expiration.
[38] It is important to note that no separation of expiration days versus non-expiration days is carried out at this stage. The thesis appendix contains figures in which expiration-days versus non-expiration days are compared for S&P 600 stocks.

4.6.1 Decimal intervals

For the S&P 600, OS experienced less clustering on round numbers.[39] The clustering was evidently larger for NOS on zero- and 10 cents, which were the two decimal places that experienced the most frequent clustering (Figure 5). The Chi-square test-statistics for the OS and NOS against the uniform distribution were 381.61 (P-VALUE = 0.000) and 427.91 (P-VALUE = 0.000), respectively. In order to test the statistical differences between the two distributions, an F-distribution test was used. The F-test is a ratio of the two Chi-square test statistics:

$$F_{diff} = \frac{Chi_O}{Chi_{NO}} \sim F_{K,N}$$

The F-statistic was 1.12 (P-VALUE = 0.2849), which indicates that the statistical difference between the two distributions was insignificant. However, although the statistical tools failed to reject the difference between OS and NOS, the economic impact may be significant. The clustering for the two groups is illustrated in Figure 5, which enhances the larger clustering for NOS. This graphical presentation gives a better overview of the data. There are 20 bins with nickels and dimes ($0.00, $0.05, $0.10 etc.) and the clustering for NOS was larger than for OS in 12 of these bins. This indicates that OS behaved more uniformly than NOS.

Both the S&P 100 and the Nasdaq 100 contain only OS.[40] The S&P 100 and the Nasdaq 100 experienced much less clustering relative to OS in the S&P 600. The Chi-square test statistics for the S&P 100 and the Nasdaq 100 were 173.28 (p-value = 0.000) and 130.54 (p-value = 0.019), respectively. Thus, the decimal clustering at the

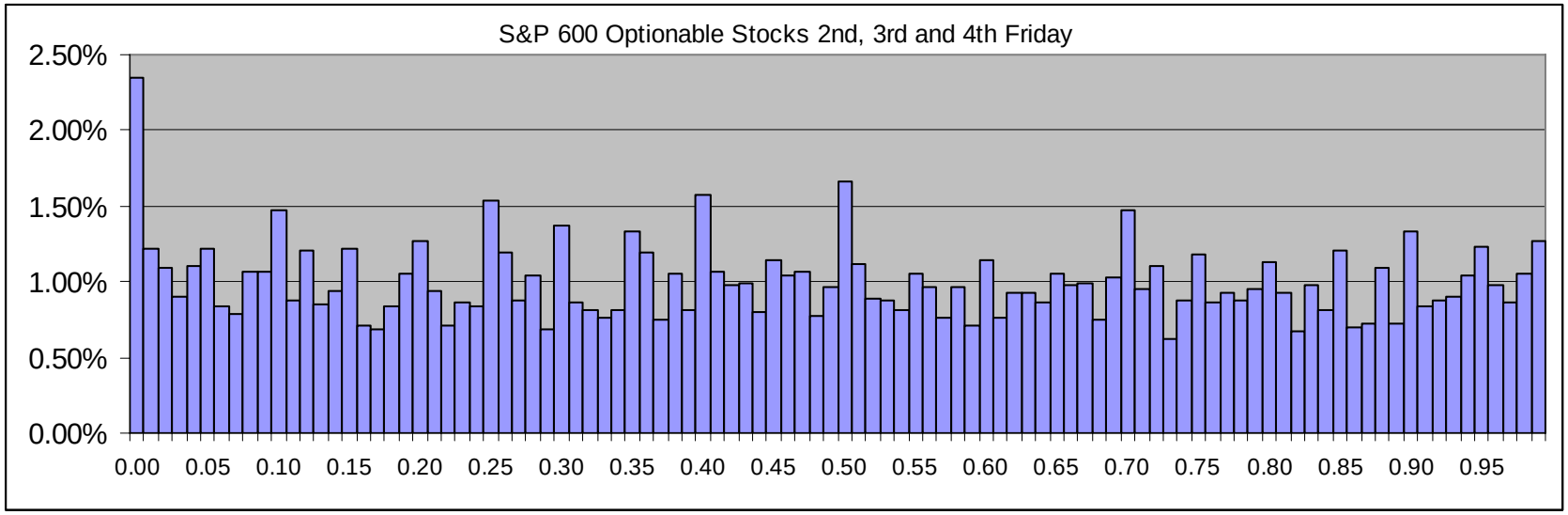
[39] The data employed in this section consist of the whole dataset and the SECOND-, THIRD- and FOURTH Friday of every month from February 2006 including August 2006 are included.
[40] That is, all the stocks in both indices are optionable, but occasionally, stocks, such as Lucent Tech. fall below $3.00 in which no new options are issued.

Nasdaq 100 could not be statistically separated from the theoretical, uniform, clustering. Figure 6 demonstrates the clustering for both the S&P 100 and the Nasdaq 100. The F-statistic that compared the S&P 100 to the Nasdaq 100 was 1.33 (p-value = 0.0803), which suggests that, at the 10%-level, the two distributions were different. Although both indices were much closer to the uniform distribution than that of the S&P 600 index, the Nasdaq 100 stocks were considerably more uniform relative to the S&P 100. A discussion to why the Nasdaq and S&P stocks differ is provided below.

Table 2: Summarises the observed clustering for all sub-sections that were studied in this chapter. It was argued that optionability and exchange trading would make the distribution of stock prices more efficient (i.e. clustering according to the uniform distribution) than non-optionability and OTC trading. For the given sample, the evidence overwhelmingly supported the hypothesis that optionability is efficiency enhancing. Across all S&P 600 observations, the OS could not be statistically differentiated from the uniform distribution, at the 5% level. The most striking difference between OS and NOS was observed among OTC stocks. In this group, the OS were indistinguishable from the uniform distribution, whereas the NOS had a p-value of 0.001 against the uniform distribution. OS were uniform (p-value = 0.0928) for listed stocks, whereas NOS rejected the uniform distribution, at the 5% level. Based on market capitalisation, the S&P 100, Nasdaq 100 and S&P 600 represent 45-, 10.6- and 3.75 percent, respectively of all United States equities. * Represents significance at the 1%-level, ** 5%-level and *** 10%-level.

	n	0.25	0.50	0.75	1.00	Pearson[41]	Prob-P	LR[42]	Prob-LR
All stocks	16,768	0.26342	0.24899	0.24445	0.24314	17.364*	0.00059	17.198*	0.00064
Replacing	16,542.68	0.25339	0.25238	0.24778	0.24645	2.290	0.51443	2.290	0.51437
EW: explained in footnote 34	16,768	0.24584	0.25485	0.25031	0.24900	2.809	0.42204	2.805	0.42262
Optionable (SP 600 only)	6,727	0.26089	0.25509	0.24260	0.24142	7.343***	0.06174	7.331***	0.06207
Non-optionable (SP 600 only)	5,873	0.27363	0.24638	0.24638	0.23361	20.037*	0.00017	19.788*	0.00019
OTC (SP 600 only)	6,081	0.27019	0.24897	0.23911	0.24174	14.485*	0.00231	14.290*	0.00254
NYSE (SP 600 only)	6,519	0.26369	0.25295	0.24927	0.23408	11.734*	0.00835	11.791*	0.00813
Optionable OTC	3,796	0.26080	0.24947	0.24552	0.24420	2.590	0.45925	2.571	0.46260
Non-optionable OTC	2,285	0.28578	0.24814	0.22845	0.23764	17.374*	0.00059	17.004*	0.00071
Optionable AMEX/NYSE	2,931	0.26100	0.262368	0.23883	0.23780	6.421***	0.09283	6.423***	0.09275
Non-optionable AMEX/NYSE	3,588	0.265886	0.245262	0.257804	0.2310479	9.973**	0.01880	10.027**	0.018338

[41] See page 22 for an explanation of the Pearson Chi-square tests.
[42] See page 22 for an explanation of the log likelihood ratio Chi-square test.



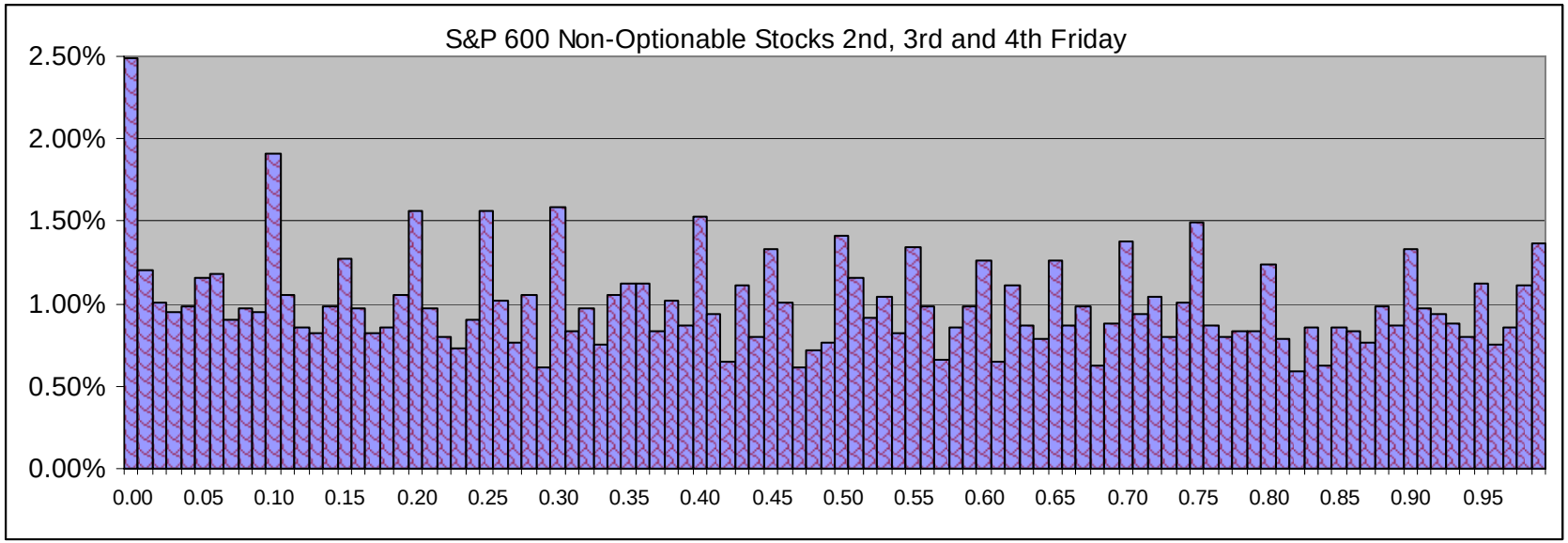
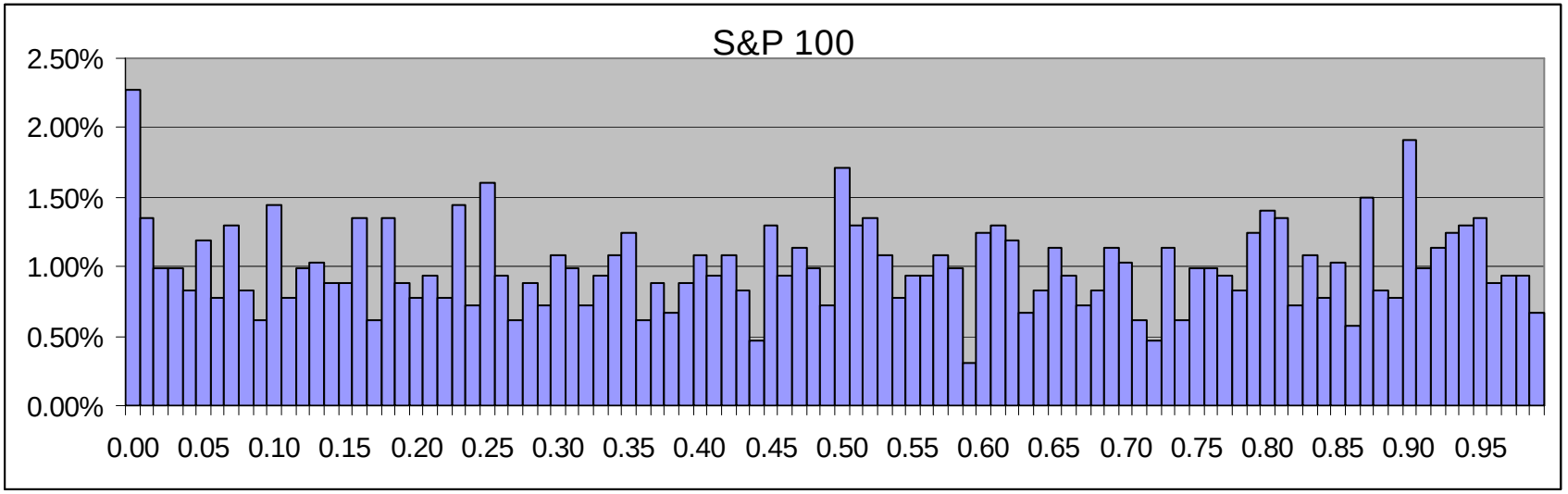


Figure 5: Illustrates the observed clustering for OS and NOS. The clustering on round numbers appears to be more prominent for the NOS, particularly in the zero- and 10-cent bin. Among the 10 nickel- and 10 dime bins, 12 bins experienced larger clustering among the NOS. The data employed in this section consist of all S&P 600 stocks and the SECOND-, THIRD- and FOURTH Friday of every month, from February 2006 including August 2006, are included. There are 6,727 OS and 5,873 NOS. A detailed analysis of S&P 600-clustering, including a comparison of expiration-day clustering and non-expiration-day clustering, is included in Appendix 4.



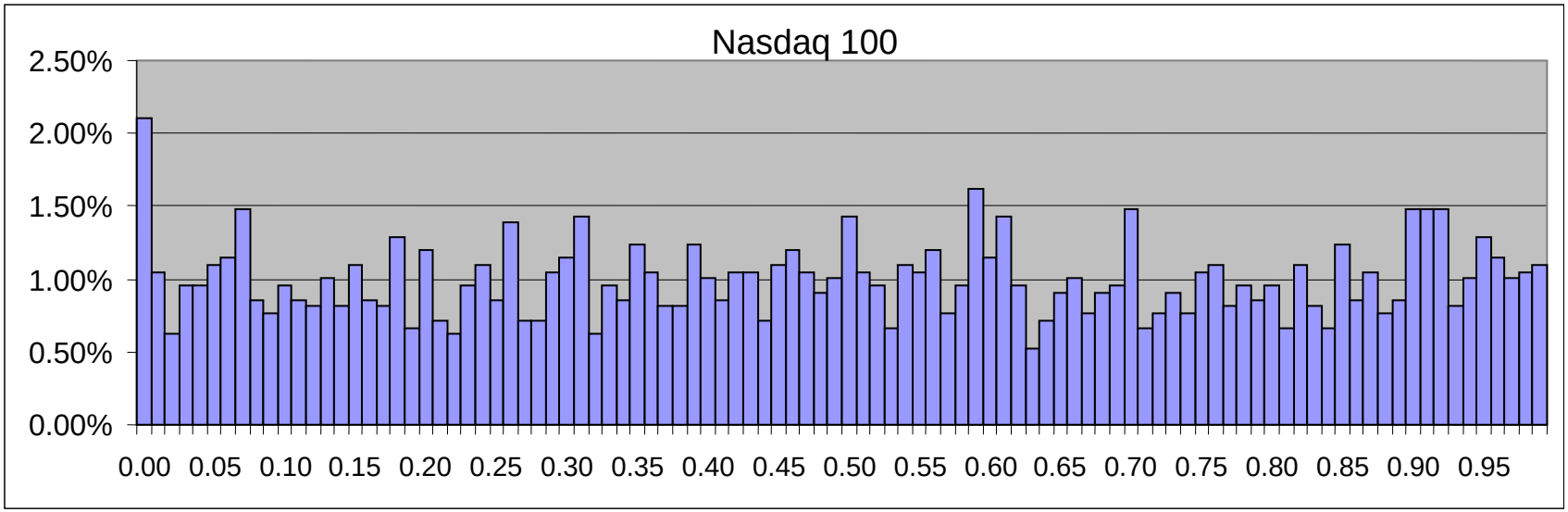


Figure 6: Illustrates clustering for stocks in the S&P 100 and the Nasdaq 100. Statistically, the clustering of Nasdaq 100 stocks was no different from the uniform distribution, at the 1% level. The S&P 100 clustering was significantly different from the uniform distribution at all levels. There were 2,092 Nasdaq observations (8 observations were missing) and 1,932 S&P 100 observations (there were 8 stocks that were included in both indices and these were removed from the S&P 100 to avoid double-counting). The SECOND-, THIRD- and FOURTH Fridays of every month from February 2006 including August 2006 were included.

4.7 The S&P 600

In order to account for differences, such as market capitalisation, volatility and volume, the S&P 600 index provides an ideal starting point. The S&P 600 index includes shares of similar characteristics.[43] However, the most desirable property of the S&P 600 is that roughly half of the stocks are optionable and the rest non-optionable. This provides an excellent opportunity to compare stocks with similar characteristics and analyse the effects of optionability on "similar" stocks. Such an analysis is impossible to perform for *mega-cap* stocks, such as those in the S&P 100 and, to a certain degree, the Nasdaq 100, because they are all optionable. The absence of non-optionable mega-cap stocks makes it difficult to draw any inferences of the effects of options on such stocks. In this section, only stocks in the S&P 600 are included in the analysis.

4.7.1 Larger intervals

The automatic exercise regime calls for an analysis of clustering on larger intervals than decimals. Twenty-five-cent ITM options are automatically exercised and the 25 cents threshold motivates the use of 25-cent intervals. The above evidence suggested that the observed data followed the uniform distribution more closely on larger intervals. In this section, OS and NOS are matched to compare the clustering on larger intervals. Optionable Stocks are expected to match the uniform

[43] According to Standard and Poor's, the criteria for being included in the S&P 600 index are: the companies must be American; the market capitalisation has to be between $300 million - $1 billion, firms have to be financially viable, which is measured by four consecutive quarters of positive earnings; firms liquidity has to be adequate, which is determined by a dollar value traded to market capitalisation-ratio of more than 0.3; firms in the index represent sectors in line with the sector balance required; and firms must be operating companies. – These criteria were retrieved on 9 October 2006 from:
http://www2.standardandpoors.com/servlet/Satellite?pagename=sp/Page/IndicesBrowseMethodologyPg&r=1&l=EN&b=4&s=6&ig=51&i=70&si=&d=&xcd=600&f=3

It is important to note that the above criteria are only the requirements for new firms to be accepted into the index. A number of S&P 600 firms' market capitalisation is well above the $1 billion ceiling.

distribution more closely than the NOS because of the increased informational efficiency brought about by options. The Chi-square test statistics for OS and NOS were 7.33 (p-value = 0.0621) and 19.788 (p-value = 0.000), respectively. This suggests that OS could not be differentiated from the uniform distribution, at the 5%-level, whereas NOS were significantly different from the uniform, at all levels. Figure 7 illustrates the frequencies across the four 25-cent intervals for both OS and NOS. Note that the OS are closer to the uniform (25%) for all but one interval.

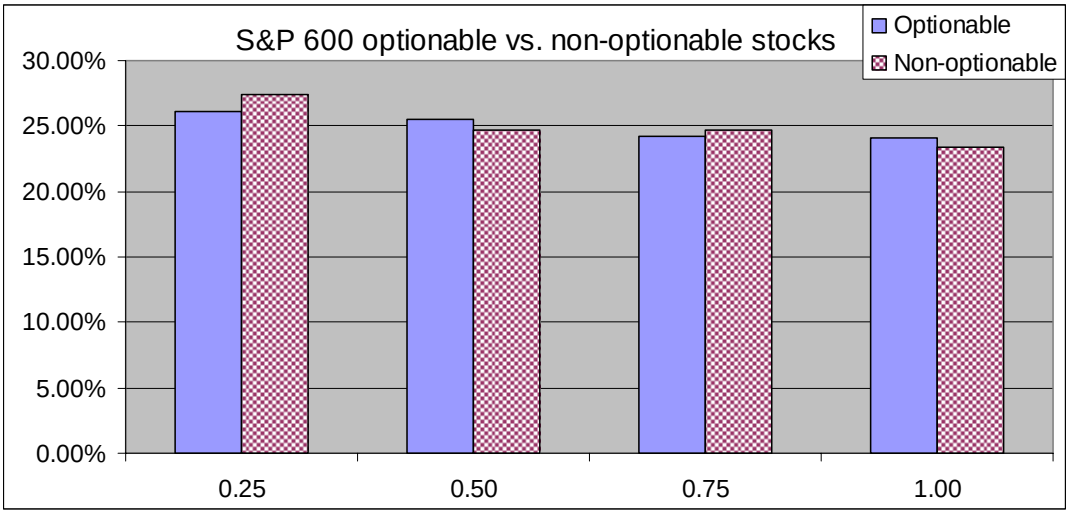


Figure 7: Illustrates the observed clustering for optionable stocks (OS) and non-optionable stocks (NOS). OS were closer to the uniform distribution in three of the four bins. Hence, OS behaved according to the theoretical distribution. Furthermore, the Chi-square test statistics for OS and NOS were 7.33 (p-value = 0.062) and 19.788 (p-value = 0.000) respectively. Hence, OS were statistically indistinguishable from the uniform distribution, at the 5%-level, whereas NOS were significantly different. The sample consist of only S&P 600 stocks, which ensured the comparison of "like with like." There are 6,727 OS and 5,873 NOS and the SECOND-, THIRD- and FOURTH Friday of every month from February 2006 including August 2006 is included. Note that the clustering and test statistics for the clustering analysis is summarised in Table 3 below. The 0.25, 0.50, 0.75, and 1.00-bars contain the [0, 24], [25, 49], [50, 74], and [75, 99]-decimals, respectively.

4.8 Over the counter vs. exchange traded stocks

There are several institutional differences between the Over-the-Counter (OTC) market and the listed market. Existing literature is in short supply of theories that explain why OTC and listed stocks experience different clustering, in particular with respect to optionability. This section considers a few microstructure-differences between the two markets.

4.8.1 Order preferencing

As shown in this study, the observed differences in clustering between OTC- and listed stocks may be attributed to different *ordering preference rules* for different markets. Ordering preference rules dictate the ranking of customers' buy and sell orders. Listed stocks (NYSE and AMEX) trade under a "specialist" system, and orders are ranked by 1) price (highest bid buys first) and 2) time (the earliest order has preference if two or more orders are similar). However, at Nasdaq orders are ranked by price only. There is no established theory of the impact of ordering preference rules on asset prices. However, it is plausible that different rules induce clustering in one system, but not in the other. For example, Bessembinder (1999) shows that order preferencing arrangements at Nasdaq lead to higher trading costs. Harris (1991) shows that investors minimise trading costs, which is achieved by clustering on round numbers because it lowers negotiation costs. Therefore, it seems reasonable that because trading costs are larger at the Nasdaq, investors have higher incentives to minimise negotiation costs by converging on round numbers (clustering) at the Nasdaq relative to NYSE.

4.8.2 The number of market makers

On the Nasdaq, there are often numerous market makers making a market in each stock. The competition between the individual market makers improve price discovery, which is a material characteristic of efficient markets. Therefore, it is reasonable to assume that the more market makers following an individual stock, the more efficient the market for this stock becomes. It also seems likely that the marginal benefit of each additional market maker diminishes, as the pool of market makers gets larger. Hence, additional market makers will improve the efficiency of stocks that do not have a very large market maker following. The S&P 600 index consists of stocks that do not have very large market maker followings. These stocks are expected to benefit from additional market makers. It is, therefore, interesting to note that the Chicago Board Options Exchange (CBOE) assigns additional market makers to any stock that becomes optionable relative to when they are non-optionable (IOSCO, 1999). Consequently, because of the increased number of market makers, all else constant, optionable stocks should be more efficient that non-optionable stocks.

4.9 Clustering of S&P 600 Stocks on 25-Cent intervals

A. Nasdaq

This section summarises the clustering analysis for 25-cent intervals. Only stocks in the S&P 600 were included in this analysis. Figure 8 illustrates the clustering for optionable and non-optionable Nasdaq (OTC) stocks. From the graph, it is obvious that the optionable stocks (OS) were distributed very close to the uniform distribution relative to the non-optionable stocks (NOS). In fact, the goodness-of-fit test did not manage to reject the hypothesis that OS are uniformly distributed. The Chi-square test statistic was 2.57 (p-value = 0.46). Hence, Nasdaq optionable stocks in the S&P 600 index were uniformly distributed for the given sample period.

Non-Optionable Stocks behaved differently from the uniform distribution, which is clearly illustrated in Figure 8. In particular, the 0 to 25-cent interval experienced much heavier clustering than the three other intervals. Against the hypothesis that NOS are uniformly distributed, the Chi-square test statistic was 17.00 (p-value = 0.000). Hence, NOS were distributed significantly different from the uniform distribution.

S&P 600 Optionable vs. Non-optionable OTC stocks.

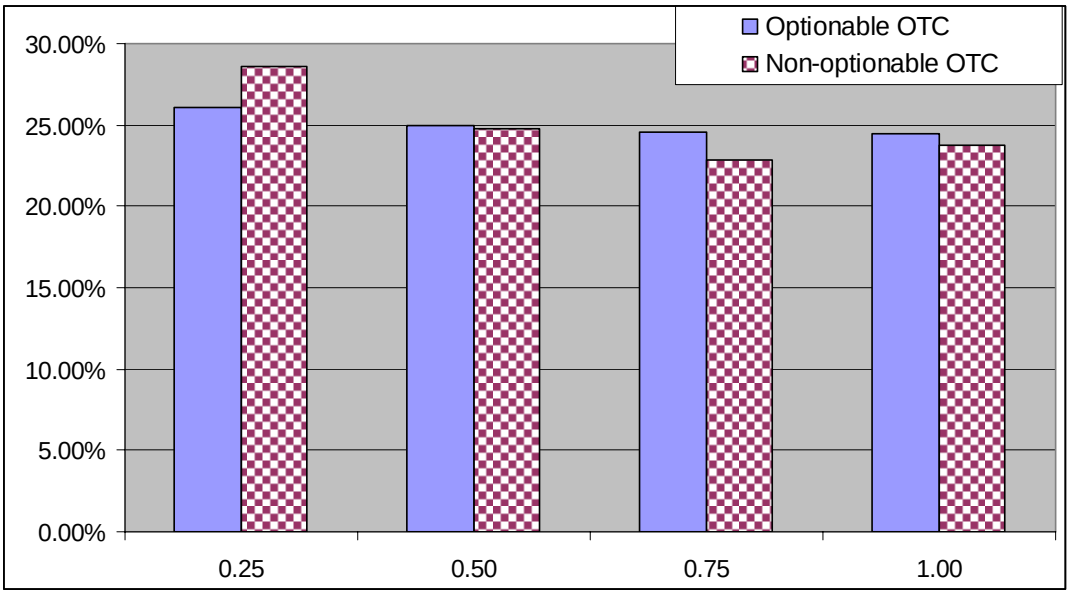


Figure 8: Illustrates the clustering for Nasdaq optionable stocks (OS) and Nasdaq non-optionable stocks (NOS) in the S&P 600. Hence, this figure employs only the Nasdaq data from Figure 7. OS were statistically indistinguishable from the uniform distribution (p-value = 0.46), whereas the NOS rejected the uniform distribution at all levels (p-value = 0.000). Note that the OS are very close to the uniform or 25% line. There were 3,796 OS OTC and 2,285 NOS OTC. The sample includes the SECOND-, THIRD- and FOURTH Friday of every month from February 2006 including August 2006. The 0.25, 0.50, 0.75, and 1.00-bars contain the [0, 24], [25, 49], [50, 74], and [75, 99]-decimals, respectively. The clustering and test statistics for this analysis is summarised in Table 3 below. It should be noted that this analysis in extended to control for differences in clustering on options expiration days. An extension of this analysis is provided in Figure 5A, panels C and D, in Appendix 5.

B. Listed Stocks: NYSE and AMEX

Figure 9 illustrates the clustering across the four 25-cent intervals for OS and NOS that were listed on the NYSE and the AMEX. Also for these stocks, OS were insignificantly different from the uniform distribution. The Chi-square test statistic of the hypothesis that OS are uniformly distributed was 6.42 (p-value = 0.093). Hence, at the 5%-level, OS were insignificantly different from the uniform distribution.

Non-Optionable Stocks were, again, significantly different from the uniform distribution at the 5%-level. The Chi-square test statistic of NOS against the uniform distribution was 10.03 (p-value = 0.018). Hence, for the NYSE/AMEX stocks in the S&P 600 index, the clustering of OS was uniformly distributed, at the

5%-level, across the given intervals, whereas NOS were significantly different from the uniform distribution, at the 5%-level.

S&P 600 Optionable vs. Non-optionable AMEX/NYSE stocks.

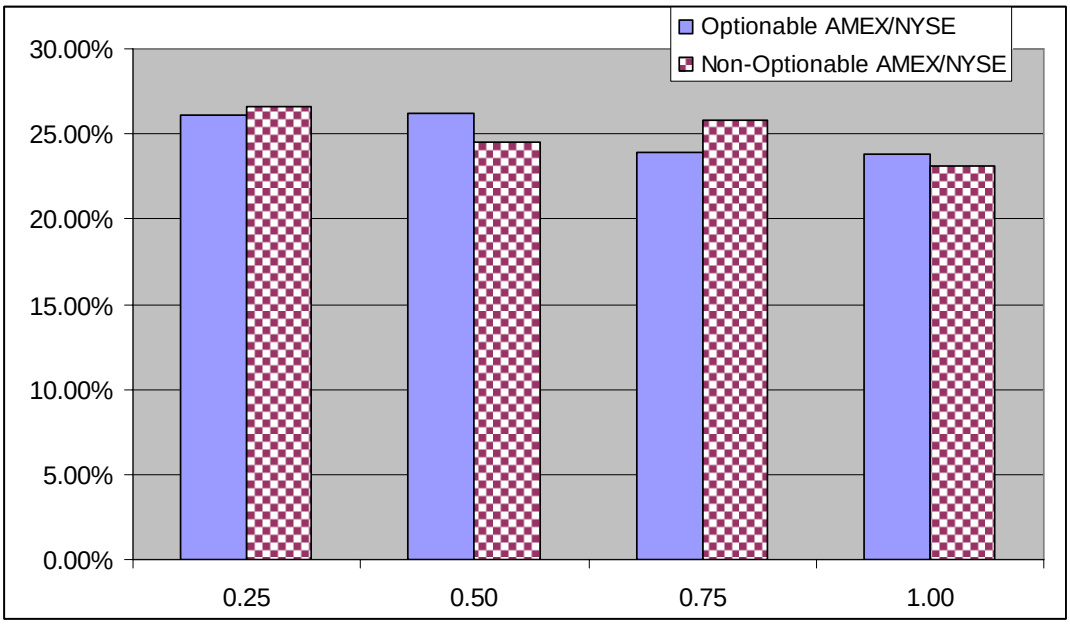


Figure 9: Illustrates clustering for listed optionable stocks (OS) and listed non-optionable stocks (NOS) in the S&P 600. OS were no different from the uniform distribution at the 5%-level (p-value = 0.093), whereas the NOS, again, were significantly different from the uniform distribution at the 5%-level (p-value = 0.019). On expiration days, there were 962 Listed OS and 1189 Listed NOS. Across all Fridays, there were 2,931 listed OS and 3,588 listed NOS. The sample includes the SECOND-, THIRD- and FOURTH Friday of every month from February 2006 including August 2006. The 0.25, 0.50, 0.75, and 1.00-bars contain the [0, 24], [25, 49], [50, 74], and [75, 99]-decimals, respectively. The clustering and test statistics for this analysis is summarised in Table 3 below. It should be noted that this analysis in extended to control for differences in clustering on options expiration days. An extension of this analysis is provided in Figure 5A, panels A and B, in Appendix 5.

4.10 Summary

Above, the work of Ikenberry et al. (2003) was replicated. The analysis confirmed their results that there was clustering on round numbers. However, the above evidence suggested that the clustering was less profound than their results, which can be explained by different data samples. The novelty of this chapter is that optionable and non-optionable stocks were analysed separately. The separation was motivated by Ni et al's (2005) arguments that optionable stocks (OS) behave differently and inefficiently relative to non-optionable stocks (NOS).

The key finding of this chapter was that OS, across Fridays before, at and after options expiration days, were distributed across the uniform distribution, whereas NOS were significantly different from the uniform distribution. This implies that OS were more efficient relative to NOS given the conservative assumption that stock prices were uniformly distributed across 25-cent intervals. Furthermore, Figure 5A, in Appendix 5, showed that listed S&P 600 stocks, OS and NOS, became increasingly uniform on expiration days. OS OTC stocks, however, experienced a shift away from the uniform distribution on expiration days, whereas NOS became increasingly uniform on expiration days.

Question 3: Is the efficiency of Optionable and Non-Optionable stocks different; if it is, is the difference more prominent on option expiration days?

"People who argue that speculation is generally destabilizing seldom realize that this is largely equivalent to saying that speculators lose money, since speculation can be destabilizing in general only if speculators on average sell when currency is low in price and buy when it is high."

- Milton Friedman (1953)

Ni et al. (2005) present evidence that option-trading cause the prices of options' underlying assets to change. However, they ignore several economically important factors in their analysis. In particular, they do not establish an expected distribution that the observed data can be compared to, nor do they pay much attention to the variation in strike price increments (strike deltas) across different stocks.[44] The lack of a theoretical clustering makes it impossible to infer whether the observed behaviour of stock prices is efficient or inefficient. In the previous (Question Two) section it is established that the distribution of stock prices is approximately uniform. Therefore, the uniform distribution is employed as the benchmark of efficiency, which is consistent with the unconditional expectation of stock prices.[45] Furthermore, the current thesis shows that sensitivity to the expected distribution and the strike deltas is crucial to measure efficiency.

Parts of Ni et al.'s (2005) analysis are replicated here, but adjustments are made to accommodate for necessary extensions of their analysis. They suggest that market manipulation and hedge rebalancing, on expiration days, cause changes to the prices of OS. This implies that OS are less efficient than what their non-optionable counterparts are. Their results are in stark contrast to existing literature that affirms

[44] Ni et al. (2005) acknowledged that strike price increments differ across different stocks, but were insensitive to this fact when analysing their data. They performed two sets of analyses where the strike price increments were either $2.50 or $5.00 for all stocks.

[45] The unconditional expectation refer to the fact the in the absence of any frictions in the marketplace or any behavioural biases, the distribution should be completely uniform.

that option trading enhances informational efficiency, which implies that OS should be closer to their theoretical values than NOS (Mayhew, 2000). If 'Friedman's proposition' (1953) carries any weight, the behaviour of speculators (firm proprietary traders in this case), as described by Ni et al. (2005), should enhance market efficiency rather than cause distortions in the marketplace.[46] However, there is a chance that underlying asset prices may be affected by their options, either negatively or positively; and the commitment of this thesis is to establish whether OS and NOS behave differently and if they do, which group is more efficient (closer to its expected value). The *a priory* expectation is that prices of OS, due to increased informational efficiency from derivative trading, should be closer to the expected uniform distribution than that of NOS.

Because Ni et al. (2005) suggest that optionable stock-prices are pulled (pushed) closer to the clustering strike price relative to NOS on expiration Fridays, the question to be asked is: Does the pull/push improve or aggravate efficiency? Another expectation of this thesis is that the expiration-day clustering of OS is closer to the theoretical clustering, relative to that of non-expiration days. This expectation is motivated by the enormous increase in trading activity on expiration-days, which fuels the flow of information and ought to enhance efficiency. Therefore, following hypotheses are put to the test:

1. Ho: The clustering of optionable stocks is closer to the expected clustering relative to that of non-optionable stocks.
2. Ho: The clustering of optionable stocks on expiration days is closer to the expected clustering relative to that of non-expiration days.

[46] 'Friedman's proposition' is the quote at the beginning of the current chapter.

4.11 Optionable vs. non-optionable stocks.

Before carrying out the tests of the two hypotheses, a brief review of the methodology is appropriate. Above, it is established that strike deltas, accompanied with the assumption that the theoretical distribution is uniform, determine the expected clustering around option strike prices. The observed and the expected clustering were calculated for each stock in the data set. To ensure robustness, both a t-test and a Z-test were employed to allow statistical discrimination between the observed and the expected clustering.[47] The two tests employ slightly different distributional assumptions in which the t-distribution has "fatter" tails relative to the normal (Z) distribution. Although the absolute test statistics produced by the Z-tests were larger than the t-tests, the rejection decisions were relatively insensitive to the distribution that the tests were based on. This is demonstrated in Table 3, Table 4, Table 5 and Table 6 below.

When the clustering around options strike prices was calculated, OS and NOS were separated. Four main groups of data were analysed: all stocks, S&P 100, Nasdaq 100 and S&P 600. The "all stocks"-group and the S&P 600 group were split into several sub-samples to account for optionability, exchange trading, expiration and different automatic exercise thresholds. However, only stocks in the S&P 600 are suitable for tests of the first hypothesis because of the absence of NOS in the S&P 100 and Nasdaq 100.

The second hypothesis was tested for all three indices. Again, the conjecture is that optionability enhances market efficiency and, consequently, the distribution of OS is expected to be closer to the uniform distribution relative to that of NOS. Table 3

[47] The two test statistics are described in detail above.

illustrates the test statistics for the tests of the hypothesis that the observed clustering around an options strike price is uniformly distributed.

4.12 Results

Across all stocks in the S&P 100, Nasdaq 100, and S&P 100, the results indicate that the clustering of OS was larger than the expected clustering (Table 3). However, the clustering of NOS was significantly less than the expected clustering. Notably, the S&P 100 stocks experienced no clustering on either of the two intervals. The clustering of Nasdaq 100 optionable stocks was significant, at the 1%-level, on the narrower interval, but insignificant on the wider interval. The S&P 600 stocks, however, experienced a small amount of clustering. Listed optionable S&P 600 stocks did marginally cluster at the narrower interval, but not on the wider interval. However, Nasdaq optionable stocks experienced excess negative clustering, at the 1%-level, on the narrow interval. Listed non-optionable S&P 600 stocks experienced excess clustering on both intervals at the 10%-level. Nasdaq non-optionable stocks, however, experiences strong negative excess clustering at the wide interval (significant at the 1%-level), but weak negative excess clustering on the narrow interval (significant at the 1%-level).

From above, it is known that the uniform distribution is a “conservative” expectation and that, normally; there is a natural clustering on options strike prices. Hence, if the true expected clustering could be calculated, the expected clustering on strike prices would increase and the distance between the observed clustering for OS would become narrower, whereas the gap between the NOS and the expected clustering would widen. It should be noted that these data represent Fridays before, at and after expiration, respectively. A more comprehensive analysis follows subsequently.

Table 3: Summarises tests of the hypothesis that the observed clustering equals the expected clustering around options strike prices. The highlighted rows represent the Z-statistics of the test that observed clustering is equal to expected clustering, whereas the two other rows contain t-statistics. Two intervals were tested. Both intervals have the clustering strike price as the midpoint, and the 25-cent interval stretches from 24 cents below the clustering strike to 24 cents above. Hence, there are 49 basis points on this interval. For the '15-cent' interval, there are 29 basis points on the interval. "Listed" means stocks that are exchange traded (NYSE and AMEX) and OTC are Nasdaq stocks. It should be noted that the number of non-optionable stocks in the S&P 100 and Nasdaq 100 is insufficiently large to draw any inferences from. * Represents significance at the 1%-level, ** 5%-level and *** 10%-level.

							S&P 600					
	All stocks		**S&P 100**		**Nasdaq 100**		OS			NOS		
	OS	NOS	OS	NOS	OS	NOS	All	Listed	OTC	All	Listed	OTC
Z 25	2.119**	-3.114*	0.254	-0.308	0.429	0.045	2.358*	NA	NA	-3.111*	NA	NA
t 25	2.037**	-3.217*	0.246	-0.319	0.414	0.042	2.259**	1.553	1.648	-3.214*	-1.691***	-3.029*
Z 15	3.926*	-2.449*	0.529	-0.793	2.274**	0.473	3.440*	NA	NA	-2.441*	NA	NA
t 15	3.693*	-2.558*	0.513	-1.080	2.112**	0.407	3.220*	1.656***	2.803*	-2.550*	-1.695***	-1.955***
Observations	10,709	5,915	1,911	21	2,071	21	6,727	2,890	3,837	5,873	3,563	2,310

4.13 The S&P 600: A comprehensive analysis

The S&P 600 index is an ideal index to study for the clustering analysis. For the given sample period, this index contained 51% and 49% listed- and OTC stocks, respectively. There were, on average, 70% OS and 30% NOS. However, on any given expiration day, several of the stocks that were optionable had no volume and, as a consequent, were treated as NOS. Hence, only 53% (6,727 stocks) of the stocks were effectively optionable and 47% (5,873 stocks) non-optionable.

4.13.1 Optionable stocks

Table 4 and Table 5 (below) provide the results from the clustering analysis for all S&P 600 stocks. It should be noted that the results in the tables are graphically depicted in Appendix A1. Table 4 lists only OS and splits the sample into listed and OTC stocks; and Fridays before; at; and after expiration, respectively.[48] Furthermore, clustering within the two clustering intervals, $[K \pm 24\,cents]$ and $[K \pm 14\,cents]$, is analysed. Evidently, listed stocks (NYSE/AMEX) experienced no excess clustering on expiration Fridays. However, there was excess clustering on Fridays prior to expiration. On the narrower $[K \pm 14\,cents]$-interval, the clustering became more significant for both listed and OTC stocks. OTC stocks exhibited significantly more clustering on expiration days relative to the expected, uniform distribution. The excess clustering on expiration days for the Optionable OTC Stocks was 3.5% (t = 3.19).

[48] Information about the number of observations for each sub-group in the analysis is provided in Table 4 and 5.

Table 4: Summarises the observed and expected clustering for optionable S&P 600 stocks. Clustering on the large- $[K \pm 24\,cents]$ and the small $[K \pm 14\,cents]$ intervals has been included. Both t-statistics and Z-statistics have been included to ensure robustness. The most noteworthy observation in this table is the clustering for Nasdaq stocks on expiration Friday. This group of stocks experienced strong clustering on expiration days, whereas listed stocks clustered less on expiration Friday than Fridays before expiration. Note that the clustering on the Friday before expiration experiences 'abnormal' clustering. *** Represents statistical significance at the 10%-level, ** at the 5%-level and * at the 1%-level. These results are illustrated graphically in Appendix A1.

S&P 600 Optionable stocks											
		Interval: +/- 24 cents					**Interval: +/- 14 cents**				
		n	observed	expected	t	Z	n	observed	expected	t	Z
All	Before expiration	2,241	15.48%	14.17%	1.70***	1.78***	2,241	9.28%	8.39%	1.45	1.53
	Expiration	2,241	16.47%	14.23%	2.81*	3.03*	2,241	11.11%	8.42%	4.02*	4.58*
	After expiration	2,245	13.76%	14.29%	-0.72	-0.72	2,245	8.37%	8.46%	-0.15	-0.15
Listed	Before expiration	961	16.55%	13.05%	2.90*	3.22*	961	9.68%	7.72%	2.04**	2.27**
	Expiration	962	13.83%	13.28%	0.48	0.50	962	9.46%	7.86%	1.68***	1.85***
	After expiration	967	12.31%	13.26%	-0.88	-0.87	967	6.93%	7.85%	-1.11	-1.06
OTC	Before expiration	1,280	14.69%	15.01%	-0.33	-0.33	1,280	8.98%	8.89%	0.12	0.12
	Expiration	1,279	18.45%	14.95%	3.19*	3.51*	1,279	12.35%	8.85%	3.79*	4.41*
	After expiration	1,278	14.87%	15.08%	-0.21	-0.21	1,278	9.47%	8.92%	0.66	0.68

4.13.2 Non-optionable stocks

Several characteristics of NOS, such as lower market maker followings for OTC stocks and less trading activity, facilitate for researchers to assume that NOS have lower informational efficiency relative to OS. Table 5 below reveals that a large difference between OS and NOS stocks exists. Particularly interesting is the result that when OS experienced a positive excess clustering; NOS experienced significant negative excess clustering. Again, listed stocks could not be differentiated from the uniform distribution, but OTC stocks were significantly different from the uniform distribution at all levels (t = - 2.59) at the $[K \pm 24\,cents]$-interval. Because of the natural clustering on the strike price becoming more dominant as the interval around the strike price narrows, the negative excess clustering on the narrower $[K \pm 14\,cents]$-interval was less severe, although still significant at the 10%-level (t = 1.957).

Table 5: Summarises the observed and expected clustering for non-optionable S&P 600 stocks. Clustering on the large- $[K \pm 24\,cents]$ and the small $[K \pm 14\,cents]$ intervals has been included. Both t-statistics and Z-statistics have been included to ensure robustness. It is evident that the listed stocks experience no excess clustering on any Friday (with one exception: Fridays after expiration on the narrow interval). OTC stocks, however, experience sever negative excess clustering on expiration Fridays. In particular clustering on the wider interval is significant at the 1%-level. *** represents statistical significance at the 10%-level, ** at the 5%-level and * at the 1%-level. Graphical representations of the results in this table are illustrated in Appendix A2.

S&P 600 NOS											
		Interval: +/- 24 cents					**Interval: +/- 14 cents**				
		n	observed	expected	t	Z	n	observed	expected	t	Z
All	Before expiration	1,959	13.78%	14.54%	-0.97	-0.96	1,959	8.22%	8.61%	-0.62	-0.61
	Expiration	1,959	12.56%	14.54%	-2.62*	-2.49*	1,959	7.55%	8.60%	-1.75***	-1.66***
	After expiration	1,955	12.99%	14.54%	-2.02**	-1.94***	1,959	7.37%	8.61%	-2.09**	-1.958***
Listed	Before expiration	1,190	13.87%	14.12%	-0.25	-0.25	1,190	7.90%	8.36%	-0.58	-0.57
	Expiration	1,189	12.87%	14.14%	-1.29	-1.25	1,189	7.82%	8.37%	-0.69	-0.68
	After expiration	1,184	12.75%	14.14%	-1.41	-1.37	1,184	7.09%	8.37%	-1.69***	-1.58
OTC	Before expiration	769	13.65%	15.20%	-1.23	-1.20	769	8.71%	9.00%	-0.28	-0.28
	Expiration	770	12.08%	15.16%	-2.59*	-2.38*	770	7.14%	8.97%	-1.957***	-1.78***
	After expiration	771	13.36%	15.17%	-1.46	-1.40	771	7.78%	8.98%	-1.23	-1.16

4.14 Triple witching

Above, it is hypothesised that the increased trading activity on expiration days increases informational efficiency to the extent that a reduction in the excess clustering of OS is expected. Triple witching is the event in which stock options, stock index options and stock index futures expire on the same day. In the literature review of this thesis, it is argued that triple witching is the expiration in which trading activity is at its largest and, consequently, increased informational efficiency ought to bring OS closer to their expected values on triple witching days. The sample contains two triple witching days, March and June, which facilitates an analysis of the witching day clustering.

Because of the desirable properties of witching days, these special expirations can be employed as the ultimate test of the hypothesis that OS are more efficient than NOS on expiration days. Throughout this study, it is argued that optionability and trading-activity should increase information flows and, consequently, lead to increased efficiency of the underlying assets. The strength of these arguments will be revealed by examining the clustering of witching days in a separate analysis. Excess clustering, among optionable stocks, on witching days is in direct conflict with the prior arguments. Therefore, it is a daunting task, at this stage in the study, to test whether these fundamental assumptions of the preceding analyses are accurate.

As Table 6 illustrates, on the $[K \pm 24\,cents]$-interval, clustering of listed optionable stocks was lower than expected, but the clustering was statistically insignificant. On the same interval, clustering of Nasdaq optionable stocks was slightly larger than expected, but the excess clustering was insignificantly different from zero. On the narrow $[K \pm 14\,cents]$-interval, the results were the same for optionable stocks.

However, whereas listed non-optionable stocks were insignificantly different from the expected clustering on both intervals, Nasdaq non-optionable stocks experienced significant negative excess clustering on both intervals.

These results are exciting because they imply that OS were, in fact, more efficient than NOS. Analysis of triple witching days suggests that increased trading activity led to lower excess clustering of OS. Hence, NOS were the securities that conflicted with their theoretical behaviour and, if anything, it is the NOS that were anomalous, not the OS.

Table 6: Shows the observed and expected clustering for OS and NOS; listed and OTC stocks for the S&P 600 at triple witching days (17 March 2006 and 16 June 2006). Only NOS OTC stocks were significantly different from the uniform distribution. These data are illustrated graphically in the appendix after the reference section. Graphical representations of the results in this table are illustrated in Appendix 2. * Represents significance at the 1%-level, ** 5%-level and *** 10%-level.

S&P 600 Triple Witching

		n	observed	expected	t	Z
Interval: +/- 24 cents						
Listed	Optionable	301	11.30%	13.17%	-1.02	-0.96
	Non-optionable	323	14.24%	14.00%	0.12	0.12
OTC	Optionable	386	17.36%	15.37%	1.02	1.08
	Non-optionable	190	10.53%	15.09%	-2.01**	-1.76***
Interval: +/- 14 cents						
Listed	Optionable	301	6.98%	7.80%	-0.55	-0.53
	Non-optionable	323	8.98%	8.29%	0.43	0.45
OTC	Optionable	386	10.10%	9.10%	0.65	0.69
	Non-optionable	190	5.26%	8.93%	-2.26**	-1.77***

4.15 Regression analysis

The previous analyses focused on relatively few dimensions, such as OS/NOS, Listed/OTC, and expiration/non-expiration. Little is said about the impact of trading-activity of options, which is measure by the open interest and the put- and call volume. The open interest of an option specifies the number of outstanding contracts of this particular option: It is the 'number of long positions' (Hull, 2003). An option's volume specifies the number of a particular option being traded on any given day (Hull, 2003). Call options are more popular, as measured by trading volume and open interest relative to put options. In the dataset employed in this analysis, the call volume is 1.36 times than the put volume; and the call open interest is 1.43 times that of the put open interest. A large number of open contracts indicate that there is a large number of investors trading the options.[49] Hence, it is reasonable to assume that the increased number of investors limits the ability of a small group of dominant investors to manipulate the underlying asset. The costs of manipulating the underlying asset will increase as the number of investors who trade a particular option increases. Therefore, a negative relationship between the open interest and clustering is expected to be detected in a regression analysis.

The option volume is less clear-cut. Normally, trading activity is expected to boost informational efficiency and, consequently, lead to a negative relationship between option volume and clustering. In contrast, Ni et al. (2005) argue that firm proprietary traders open new positions (i.e. they write options) within one week of expiration and then manipulate the underlying asset so that options expire OTM. Ni et al.'s (2005) projection implies a positive relationship between volume and clustering. However, it

[49] Alternatively, it could be that a relatively small number of investors hold large blocks of options, but this is assumed an infrequent occurrence.

is important to note that it is extremely difficult to establish cause-and-effect in this analysis. For example, on expiration day, it seems plausible that investors trade options and underlying assets if the underlying price is close to the strike price. This happens because options that expire ATM or OTM are worthless after expiration. Therefore, in cases in which the stock price trades close to the strike price, the probability of the options expiring OTM increases and the investors have incentives to close out their positions before their options expire. Hence, without an established theory to explain cause-and-effect, it seems equally plausible that clustering is caused by increases in volume as well as the increased volume being caused by a stock trading near its strike price.

4.15.1 Logistic Regression

A logistic regression of clustering with the options data can help identify the correlation of trading (volume) and open interest upon clustering.[50] In addition to volume and open interest, two additional variables are included in the regressions. First, stock prices are likely to carry information about clustering. Stocks are often listed at relatively low prices in order to attract investors who are not willing to buy expensive shares. Therefore, an expensive stock has often experienced large growth in order to reach its high price. Stocks with large historical growth experience healthy interest from investors and intense scrutiny from analyst, which ought to improve efficiency of the respective stocks. Therefore, a high-priced stock, ceteris paribus, should be more efficient, and an inverse relationship between price and clustering is

[50] A brief discussion of statistical assumptions in relation to the Logistic regression is provided in Appendix 3.

expected.[51] Furthermore, the strike deltas become wider as prices increase, which causes the probability of clustering to decrease. Secondly, a control variable (OTC dummy: 1 if OTC and 0 otherwise), to account for differences between listed- and OTC stocks, is included.

Because of limited access to options data, the regression analysis is restricted to expiration days only. Furthermore, because of their desirable characteristics (discussed above), only stocks in the S&P 600 index are included in the analysis. Two regression equations are employed: one is a Logit regression where the dependent variable is a discrete zero-one clustering dummy (1 if clustering and 0 otherwise); and the other is an OLS regression in which dependent variable is the absolute log-ratio of a stock closing price divided by the clustering strike price. Although the two regressions are not testing the exact same hypotheses, the use of two equations is helping to ensure that the results are robust.

The specification of the Logit equation is:

$$C = \alpha + \beta_1 \log(CVol) + \beta_2 \log(COi) + \beta_3 \log(PVol) + \beta_4 \log(POi) + \beta_5 \log(\mathrm{P}) + \beta_6 OTC + \varepsilon$$

The OLS equation:

$$\log\left|\frac{S_T}{K}\right| = \alpha + \beta_1 \log(CVol) + \beta_2 \log(COi) + \beta_3 \log(PVol) + \beta_4 \log(POi) + \beta_5 \log(\mathrm{P}) + \beta_6 OTC + \varepsilon$$

[51] The fact that a stock has become "expensive" implies that it has grown to a high price. Stocks grow when interest in them is high. Stocks that experience growth attract attention and, consequently, such stocks ought be more informational efficient relative to stocks that are not growing to high prices. This analysis is insensitive to stock splits, which often occur when stocks reach a certain threshold.

4.15.2 OLS Regression

The dependent variable of the OLS regression is the absolute log-ratio of the official closing price to the clustering strike price. Positive regressor-coefficients indicate that the difference between the closing price, S_t, and the clustering strike price, K, is becoming larger given a unit-increase of the regressor. The wider the gap between S and K, the less clustering. In contrast, negative signs on the coefficients indicate that the larger these variables become, at the margin, the more clustering (the gap narrows). The marginal effects of the regressions are listed in Table 7.

Table 7
Regressions to Explain Clustering

This table illustrates the marginal effects of the three regressions that were carried out. Two Logit regressions, 24 and 14, were performed in which the dependent variable was a dummy variable that was equal to one if a stock's closing price was within 25 or 15 cents from an option strike price, respectively. The log(S/K) equation was estimated using an OLS regression. Significance at the 1% and 10%-levels are indicated by * and ***, respectively. NOTE: All dependent variables, except OTC, are logged. Standard errors are reported in brackets.

Marginal effects table

	24	14	log(S/K)
constant	NA	NA	0.111098*
			(0.00451)
call vol	0.02917*	0.02632*	-0.00205*
	(0.00666)	(0.00556)	(0.000455)
call OI	-0.02643*	-0.02053*	0.00167*
	(0.00788)	(0.00663)	(0.00055)
put vol	0.02785*	0.01551*	-0.00139*
	(0.00679)	(0.00561)	(0.00047)
put OI	-0.00329	0.00021	-0.000021
	(0.00781)	(0.00655)	(0.00054)
price	-0.07209*	-0.0386*	-0.02259*
	(0.01514)	(0.01265)	(0.00114)
otc	0.03477***	0.02337	-0.00089
	(0.01829)	(0.01531)	(0.00130)

As expected for the Logistic regressions, the open interest, for calls, was negatively correlated with clustering, which implies that the more outstanding options contracts, the more efficient the market was, which caused lower clustering. The put open interest had no impact on the probability of clustering.

Both the call- and put volume were positively related to the clustering, which indicates that a unit increase in volume increased the probability of clustering. It should be noted, however, that cause-and-effect is impossible to establish and that it is equally likely that the high volume is a symptom of clustering as opposed to the cause of the clustering. Because the table illustrates marginal effects (not regression coefficients), it implies that, at the margin, a unit-increase (100 contracts) in volume increased the probability of clustering on the $[K \pm 24\,cents]$-interval by 2.9%. The largest marginal effect was found for the log-price variable. There is a strong inverse relationship between clustering and price, which might suggest that the larger the stock price becomes, the more information is contained within the price and the less clustering occurs. In fact, at the margin, a dollar increase in price reduced the probability of clustering at the $[K \pm 24\,cents]$-interval by 7.2%. An additional explanation for the inverse relationship between stock price and clustering is that the strike deltas increase as the stock price increases, which reduce the probability of clustering. The OTC dummy was only significant at the 10%-level at the $[K \pm 24\,cents]$-interval, but insignificant at the $[K \pm 14\,cents]$-interval. Hence, whether a stock was listed or OTC did not affect the probability of clustering at the narrower interval.

The OLS regression was intended as a robustness check of the Logit equation. The dependent variable is, as discussed above, the log-ratio of the stock price, divided by

the strike price. Hence, in this equation, positive coefficients indicate less clustering and negative coefficients indicated more clustering. Therefore, the signs on the coefficients are expected to be opposite from those in the Logit regressions. This was true for all significant variables apart from the “price”-variable, which had the same sign in both equations. Therefore, larger priced stocks were likely to have narrower S/K-ratios. This result, however, does not necessarily compromise the results from the Logit regressions. The dependent variable in this regression is widely different from the dependent Logit-variable and the only resemblance between the two is that log(S/K)-ratios that are close to zero would qualify as clustering. It is explained above that the strike price, K, increases as the price, S, increases. Therefore, the denominator will increase as price increases, which can cause the ratio to narrow as prices increase. In addition, it is argued above that clustering on round numbers increases with price. Hence, because option strike prices normally are on round numbers, the (S/K)-ratio tend to decrease when S becomes large.

4.16 Summary

This chapter extended the preceding analyses to examine clustering of stock prices at or near options strike prices. In particular, Ni et al.'s (2005) analysis was extended to include important features, such as the employment of individual stocks' strike deltas to construct an expected clustering.

The most striking results showed that listed optionable S&P 600 stocks did not cluster on expiration days. However, optionable Nasdaq stocks did experience excess clustering on expiration Fridays, relative to non-expiration Fridays. In contrast, non-optionable Nasdaq S&P 600 stocks experienced strong negative excess clustering on expiration Fridays, whereas non-optionable listed stocks experienced no excess clustering. Furthermore, an analysis of clustering on triple witching days proved that increased informational efficiency on these days led to the clustering of optionable stocks becoming increasingly uniform, whereas non-optionable Nasdaq stocks were deviating further away from the theoretical clustering than non-witching days.

At last, a regression analysis was performed, with the purpose of establishing correlations between options market variables, such as open interest and volume, and clustering. The regressions revealed a significant negative relationship between clustering and call open interest, which indicates that the probability of clustering was reduced when the call open interest increased. No relationship between the put open interest and clustering was found. However, positive relationships between clustering and put- and call volume were found. Although cause-and-effect was not established, these results imply that increases in volume increased the probability of clustering.

Chapter 5.0: Conclusions

Question One: Sensitivity to underlying market microstructure variables, such as strike price increments and volume on options contracts, was shown to be extremely important when calculating the expected clustering of individual stocks, and determining whether stocks should be classified as optionable or non-optionable.

Question Two: Ikenberry et al. (2003) showed that stock prices are not uniformly distributed across each possible decimal-bin. However, this research has shown that uniformity across each decimal-bin is not necessary in order to employ uniformity as the theoretical distribution. In particular, the automatic exercise thresholds for In-The-Money options determined the bin-sizes. Thorough tests of uniformity across the [0, 24]-, [25, 49]-, [50, 74]-, and [75, 99]-cent intervals were performed. Particularly noteworthy were the results that optionable stocks were distributed more uniformly than non-optionable stocks, which implied that optionable stocks were more efficient than non-optionable stocks given the uniformity paradigm.

Question Three: Across all expiration Fridays, listed Optionable Stocks were insignificantly different from the uniform distribution, whereas Optionable OTC Stocks did experience some excess clustering. However, the real test of expiration-day efficiency was carried out by using Triple Witching days only. Triple Witching day clustering for Optionable Stocks was no different from the uniform distribution. However, Non-Optionable Nasdaq stocks experienced significant clustering on both intervals.

5.1 Was the objective achieved?

The main objective was to identify differences in behaviour of optionable and non-optionable stock prices on expiration days. Based on the analyses in Chapter 4, coupled with the assumption that stock prices are uniformly distributed across given intervals, it is clear that the efficiency of optionable stocks improves on option expiration days.

5.2 Limitations and key assumptions

Several assumptions are made throughout this thesis, with the most significant assumption being that stock prices are uniformly distributed across all possible bins (decimals). In a marketplace with perfect information, no transaction costs and absent behavioural/cognitive biases, the distribution should be uniform. Tests of this assumption are provided in the preceding chapter (Question Two).

Another assumption of this research is that options strike-price increments (strike deltas) are constant between Fridays before expiration, at expiration and after expiration, respectively. Because of limited access to options data, strike prices were updated every expiration-day. Consequently, strike deltas on Fridays before and after expiration were assumed equal to expiration-day strike deltas. The only exception to this rule was given to stocks that experienced a price-change that moved the price from one price-range into another price-range in which the strike deltas normally change. Generally, strike deltas are \$2.50, \$5.00 and \$10.00 for stock prices on the intervals [\$2.50, \$25.00), [\$25.00, \$200.00) and [\$200.00, ∞), respectively.[52] Although these are the general strike deltas, strike deltas are often adjusted for stock

[52] It should be noted that Ni et al.'s (2005) *JFE*-article misrepresented the strike deltas (Catania and Maberly, 2006). Their footnote 14 states 'Exercise prices below \$20 include odd integer multiples of \$2.50. Occasionally exercise prices that are not integer multiples of \$2.50 also occur…'

splits, stock dividends and large dividends.[53] Furthermore, several pilot programmes exist in which selected stocks' strike deltas are $2.50 for stock prices up to $50.00 and $70.00, respectively. Random tests of this assumption were performed and after adjusting the strike deltas of stocks that moved across the above intervals, no inconsistencies in strike deltas were detected.

5.3 Further research

There are opportunities for additional research in several directions. In this study, existing knowledge of options market microstructure was applied in a seemingly original context. However, more research into market microstructure is needed to better inform researchers who are doing options market research. Ignorance of the underlying functioning of the relevant markets can lead to flawed conclusions.

Throughout this study, the underlying paradigm was that the distribution of stock prices was uniform. It was acknowledged that this paradigm understated the expected clustering. Deriving a more precise measure of the expected clustering is a task for researchers with strong microstructure expertise and technical capabilities. The integrity of such research, however, is constrained by data reliability, and scholars or professionals with access to high-quality data would find it worthwhile to replicate and extend the above analyses.

[53] Adjusted strike prices/deltas often appear to be a non-round number. For example, a stock that is trading at $113 (closest strike = $115) before the 3:1 split will trade at $37.67 and the closest strike will be $38.33 (115/3). Consequently, the strike delta after the split will be $1.67 ($38.33-$36.67), as opposed to $5 before the split.

References

Avellaneda, M., & Lipkin, M. D. (2003). A market-induced mechanism for stock pinning. *Quantitative Finance, 3*, 417-425.

Ball, C.A., Torous, W. N., Tschoegl, A.E. (1985). The degree of price resolution: The case of the gold market. *The Journal of Futures Markets*, *5(1)*, 29-43.

Barclay, M. J., Hendershott, T., Jones, C. M., (March, 2006), "Order consolidation, price efficiency, and extreme liquidity shocks", *Simon School, University of Rochester*, Research Paper Series, Forthcoming.

Beechey, M., Gruen, D., Vickery, J. (2001). The efficient market hypothesis: A survey. *Reserve Bank of Australia*, Research discussion paper.

Bessembinder, H. (1999). Trade execution costs on Nasdaq and the NYSE: A post-reform comparison. *Journal of Financial and Quantitative Analysis*, *30(3)*, 387-407.

Black, F., Scholes, M. (1973). The pricing of options and corporate liabilities. *The Journal of Political Economy*, *81(3)*, 637-654.

Keoun, B. (2006). E*Trade, TC Ameritrade are targeted in online brokerage fraud. *Bloomberg*. Retrieved 10 November 2006 from http://www.bloomberg.com/apps/news?pid=20601103&sid=a0tTA2Ko30As&refer=us

Brav, A., Heaton, J.B. (2002). Competing theories of financial anomalies. *The Review of Financial Studies*, *15(2)*, 575-606.

Brockman, P., Michayluk, D. (1998). The persistent holiday effect: additional evidence. *Applied Economic Letters*, *5*, 205-209.

Catania, P., Maberly, E.D. (2006). The official closing versus last trade price for Nasdaq stocks: Implications for empirical research and automatic equity option exercise. *SSRN. Working paper*.

Catania , P., Maberly, E.D., Pierce, R., (2006). History of options price strikes. University of Canterbury. *Working paper*.

Chung, H., Chiang, S. (2006). Price clustering in E-mini and floor-traded index futures. *The Journal of Futures Markets*, *26(3)*, 269-295.

Cox, J., Ross, S., Rubinstein, M. (1979). Option pricing: A simplified approach. *Journal of Financial Economics*, *7*, 229-264.

De Long, J.B., Lang, K. (1989). Are all economic hypotheses false? *Journal of Political Economy*, *100(6)*, 1257-1272.

Easley, D., O'Hara, M., Srinivas, P. S. (1998). Option volume and stock prices: Evidence on where informed traders trade. *The Journal of Finance*, *53(2)*, 431-465.

Economist T. (2006). Dismal science, dismal sentence. *The Economist*, *380(8494)*, London: Sep 9, 2006, p. 90.

Fama, E. F. (1965). The behaviour of stock-market prices, *The Journal of Business*, *38(1)*, 34-106.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, *25(2)*, 383-417.

Fama, E. F. (1998). Market efficiency, long-term returns, and behavioural finance. *Journal of Financial Economics*, *49*, 283-306.

Figlewski, S., Webb, G. P. (1993). Options, short sales, and market completeness. *The Journal of Finance*, *48(2)*, 761-777.

Friedman, M. (1953), The case for flexible exchange rates. *Essays on Positive Economics*, University of Chicago Press: Chicago.

Fuhr, D. (2001). Exchange-traded funds: A primer. *Journal of Asset Management*, *2(3)*, 260-273.

Gann, W. D. (1930). *Truth of the stock tape*. Financial Guardian Publishing Co, New York. 63-64.

Glosten, L. R. (1989). Insider trading, liquidity, and the role of the monopolist specialist. *The Journal of Business*, *62(2)*, 211-235.

Granger, C. W. J., Morgenstern, O. (1970). *Predictability of stock market prices*. Heath Lexington Books, Lexington: MA.

Grossman, S. J., Miller, M. H., Cone, K. R., Fischel, D. R., Ross, D. J. (1997). Clustering and competition in asset markets. *Journal of Law and Economics*, *40(1)*, 23-60.

Harris, L. (1991). Stock price clustering and discreteness. *The Review of Financial Studies*, *4(3)*, 389-415.

Harris, L. (2003). *Trading and exchanges: Market microstructure for practitioners*. Oxford university press: New York: NY.

Herbst, A. F., Maberly, E. D. (1990). Stock index futures, expiration day volatility, and the "special" Friday opening: A note. *The Journal of Futures Markets*, *10(3)*, 323-325.

Hull, J. (2003). *Options, futures, and other derivatives*. Pearson Education Inc., Upper Saddle: NJ, 5th Ed.

Ikenberry, D., Weston J. P. (2003). Clustering in U.S. stock prices after decimalization. *Rice University*, *Working paper*.

International Organization of Securities Commissions. (1999). *The influence of market makers in the creation of liquidity*. Retrieved November 4, 2006, from http://www.iosco.org/library/

Jaffe, J., Westerfield, R. (1985). The week-end effect in common stock returns: The international evidence. *The Journal of Finance*, *40(2)*, 433-454.

Jarrow, R. A. (1994). Derivative security markets, market manipulation, and option pricing theory. *The Journal of Financial and Quantitative Analysis*, *29(2)*, 241-261.

Jessop, A. (1995). *Informed assessments: an introduction to information, entropy, statistics / Alan Jessop*. Ellis Horwood Limited., Hertfordshire.

Kahneman, D., Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47(2)*, 263-292.

Kahneman, D., Slovic, P., Tversky, A. (1982). Judgment under uncertainty: Heuristics and biases. *Cambridge University Press*: Cambridge.

Krishnan, H., Nelken, I. (2001). The effect of stock pinning upon option prices, *Risk*: December, 17-20.

Kumar, R., Sarin, A., Shastri, K. (1998). The impact of options trading on the market quality of the underlying security: An empirical analysis. *The Journal of Finance*, *53(2)*, 717-732.

MacKinlay, A.C. (1997). Event studies in economics and finance. *Journal of Economic Literature*, *35(1)*, 13-39.

Malkiel, B. G. (2003). The efficient market hypothesis and its critics. *Journal of Economic Perspectives*, *17(1)*, 59-82.

Mayhew, S. (2000). The impact of derivatives on cash markets: what have we learned?. University of Georgia, *Unpublished working paper*.

Merton, R. C. (1973). Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, *4(1)*, 141-183.

Miller, M. H. (1995). Do we really need more regulation of financial derivatives?. *Pacific-Basin Finance Journal*, *3*, 147-158.

Mitchell, J. (2001). Clustering and psychological barriers: The importance of numbers. *Journal of Futures Markets*, *21(5)*, 395-428.

Ni, S. X., Pearson, N. D., Poteshman, A. M. (2005). Stock price clustering on option expiration dates, *Journal of Financial Economics*, *78(1)*, 49-87.

Neumann, J. V., Morgenstern, O. (1944).*The theory of games and economic behaviour*. Princeton University Press: Princeton, 1st Ed.

Niederhoffer, V. (1966). A new look at clustering of stock prices. *The journal of business*, *39(2)*, 309-313.

Odean, T. (1998). Are investors reluctant to realize their losses?. *The Journal of Finance*, *53(5)*, 1775-1798.

O'Hara, M. (1995). *Market microstructure theory*. Blackwell Publishers: Cambridge, MA, 1st Ed.

Panayides, M. A. (February 7, 2004). The specialist's participation in quoted prices and the NYSE's price continuity rule. *Yale ICF Working Paper No. 04-05*, Available at SSRN: http://ssrn.com/abstract=492423

Ross, S. R. (1976). Options and efficiency. *The Quarterly Journal of Economics*, *90(1)*, 75-89.

Schwartz, A. L., Van Ness, B. F., Van Ness, R. A. (2004). Clustering in the futures market: Evidence from S&P 500 futures contracts. *The Journal of Futures Markets*, *24(5)*, 413-428.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, *19(3)*, 425-442.

Shleifer, A. (December 28, 2000). Are markets efficient? No, arbitrage is inherently risky. *The Wall Street Journal*, Eastern edition: New York, N.Y. Page 10.

Sonnemans, J. (2006). Price clustering and natural resistance points in the Dutch stock market: a natural experiment. *European Economic Review, 50(8)*, 1937-1950.

Stoll, H. R. and Whaley R. E. (1991). Expiration-day effects: What has changed?. *Financial Analyst Journal*, *47(1)*, 58-72.

Summers, L. (1986). Does the stock market rationally reflect fundamental value?. *The Journal of Finance*, *41(3)*, 591-601.

Thesis Appendices

Appendix 1: Clustering on options strike prices

A1: Optionable stocks

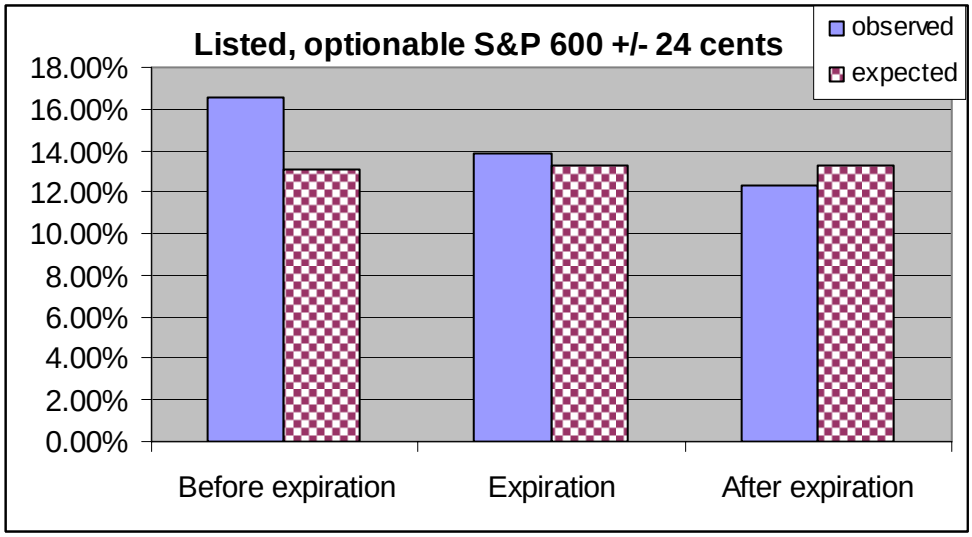


Figure 1A: Corresponds to Table 4. This figure illustrates the clustering of listed OS on the $[K \pm 24\,cents]$-interval. It is obvious that there was excess clustering on Fridays prior to expiration. Neither the expiration-day clustering nor the after-expiration-day clustering was significantly different from the uniform distribution. However, the before-expiration clustering was significantly different from the uniform distribution, (t = 2.90 and Z = 3.22).

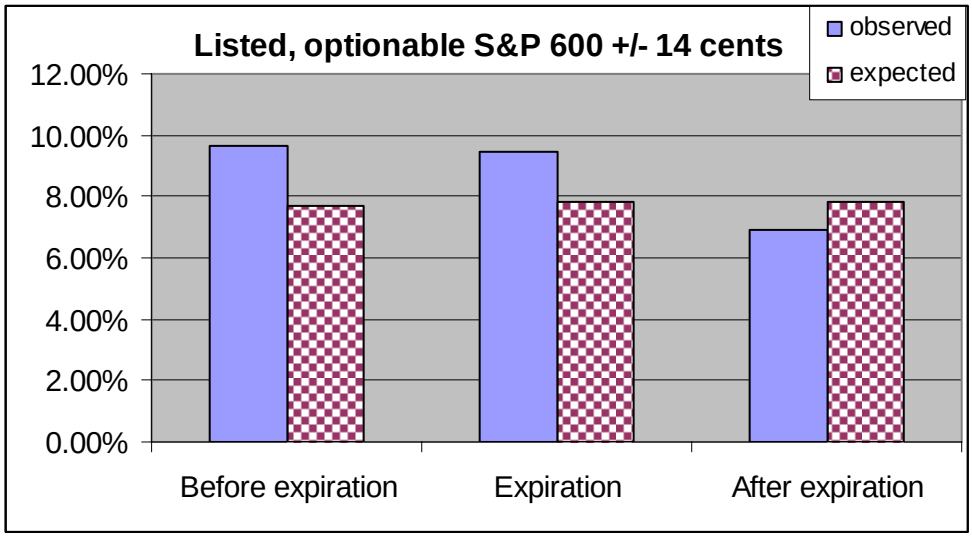


Figure 1B : Illustrates the clustering for listed OS on the $[K \pm 14\,cents]$-interval. On this narrower interval, the clustering on expiration days became significant at the 10%-level, in which the t-test was 1.68 and the Z-test was 1.85 for Expiration Friday clustering. The before-expiration clustering was significant at the 5%-level, in which the t- and Z-tests were 2.04 and 2.17, respectively.

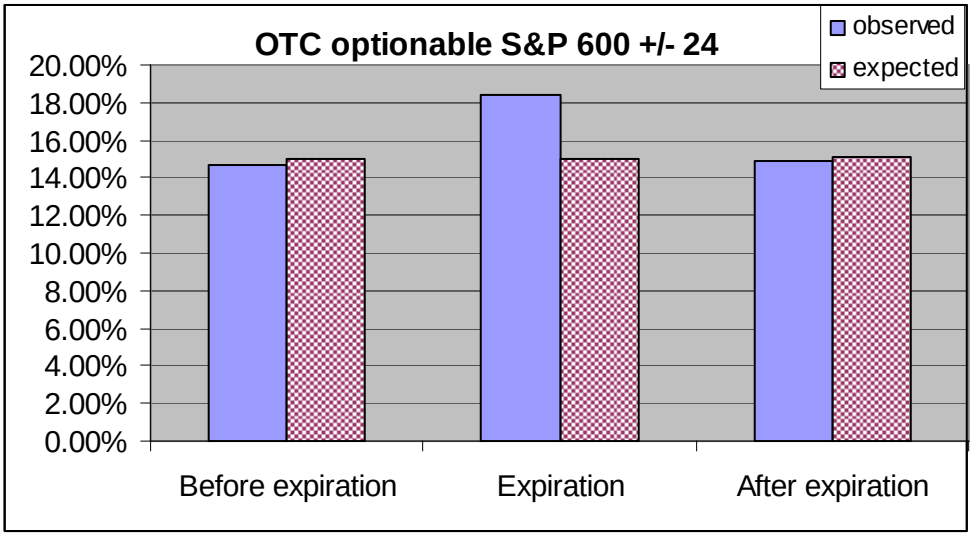


Figure 1C: Corresponds to Table 4. This figure illustrates the clustering of OTC OS on the $[K \pm 24\,cents]$-interval. It is obvious that there was excess clustering on expiration Fridays. Neither the before-expiration-day clustering, nor the after-expiration-day clustering was significantly different from the uniform distribution. However, the expiration-day clustering was significantly different from the uniform distribution, in which the t-test was 3.19 and the Z-test was 3.51.

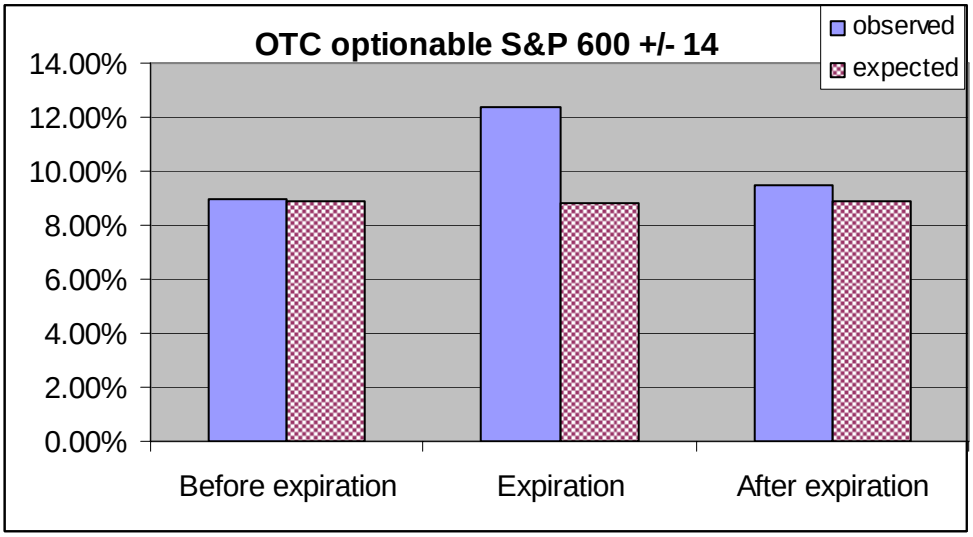


Figure 1D: Illustrates the clustering for NOS on the $[K \pm 14\,cents]$-interval. On this narrower interval, the clustering on expiration days was still significant at the 1%-level, in which the t-test was 3.79 and the Z-test was 4.41. The before-expiration-day clustering and after-expiration-day clustering were insignificant at all levels.

A2: Non-optionable stocks

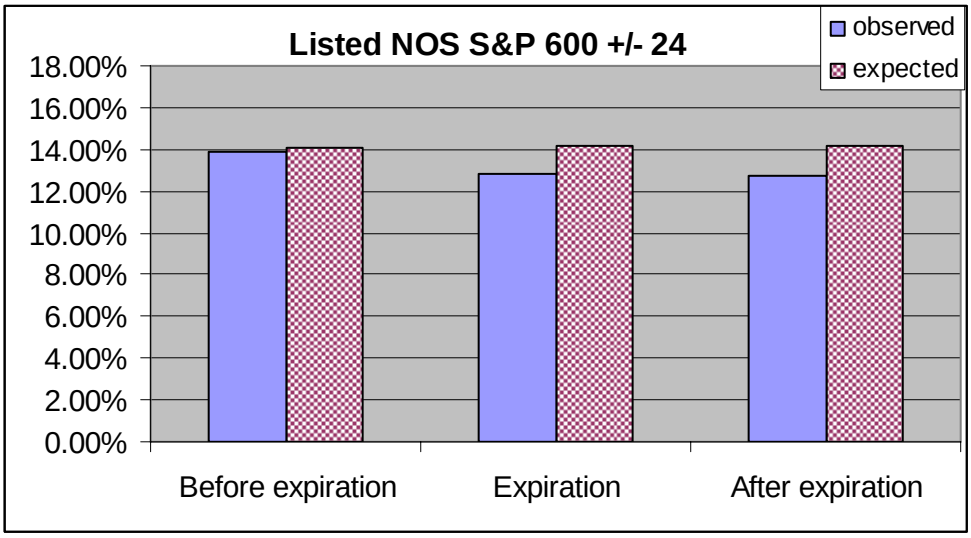


Figure 1E: Corresponds to Table 5. It illustrates clustering for the listed NOS on the $[K \pm 24\,cents]$ -interval. Statistically, there was no excess clustering for any Friday.

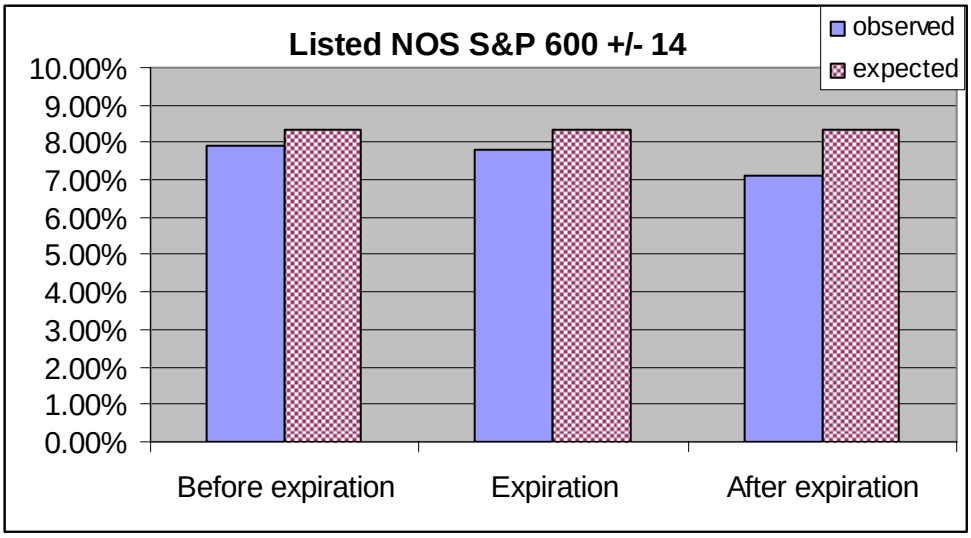


Figure 1F: Illustrates the clustering for listed NOS on the $[K \pm 14\,cents]$-interval. Only the after-expiration-day clustering was significant at the 10%-level according to the t-distribution (t = -1.69), however, the Z-test was insignificant. This was one of the rare cases in which the t- and the Z-tests were inconsistent in their rejection decision.

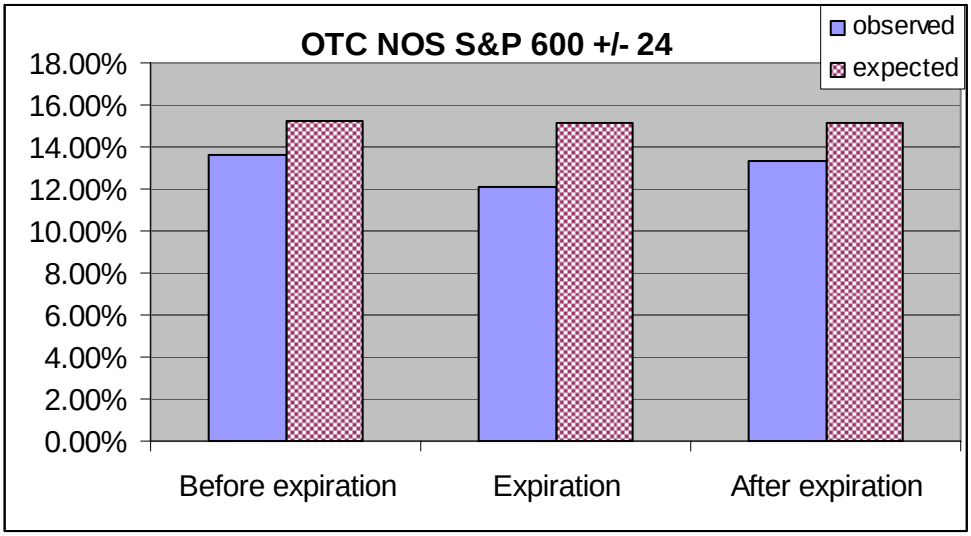


Figure 1G: Corresponds to Table 5. It illustrates the clustering for OTC NOS on the $[K \pm 24\,cents]$-interval. Negative excess clustering was significant for expiration days, at the 1%-level. The respective t- and Z-statistics were -2.59 and -2.38.

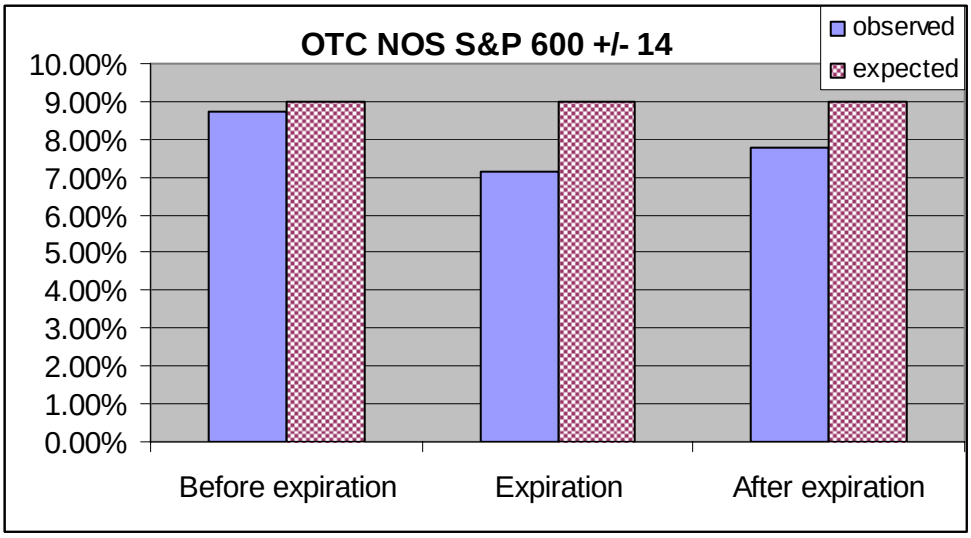


Figure 1H: Illustrates clustering of OTC NOS on the $[K \pm 14\,cents]$-interval. Here too, the negative excess clustering became, statistically, less significant on the narrower interval. The respective t- and Z-statistics were -1.957 and -1.78, which imply statistical significance at the 10%-level.

Appendix 2: Triple Witching

The following graphs illustrate stock price clustering within given intervals,

$[K \pm 14\,cents]$ and $[K \pm 24\,cents]$, of an option strike price.

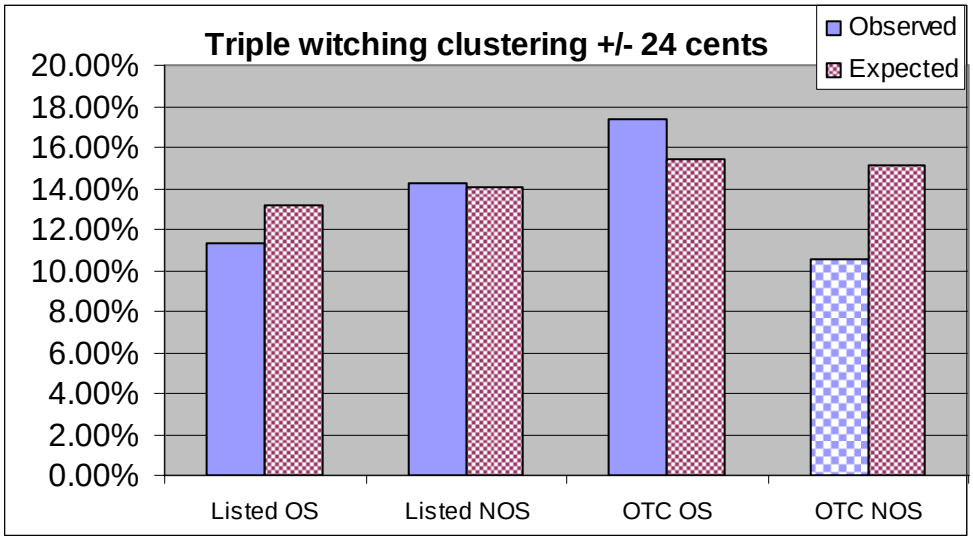


Figure 2A: Illustrates the triple witching clustering for listed OS and NOS, and OTC OS and NOS at the $[K \pm 24\,cents]$-interval. All OS were insignificantly different from the uniform distribution, whereas the OTC NOS experienced negative excess clustering (t = -2.01).

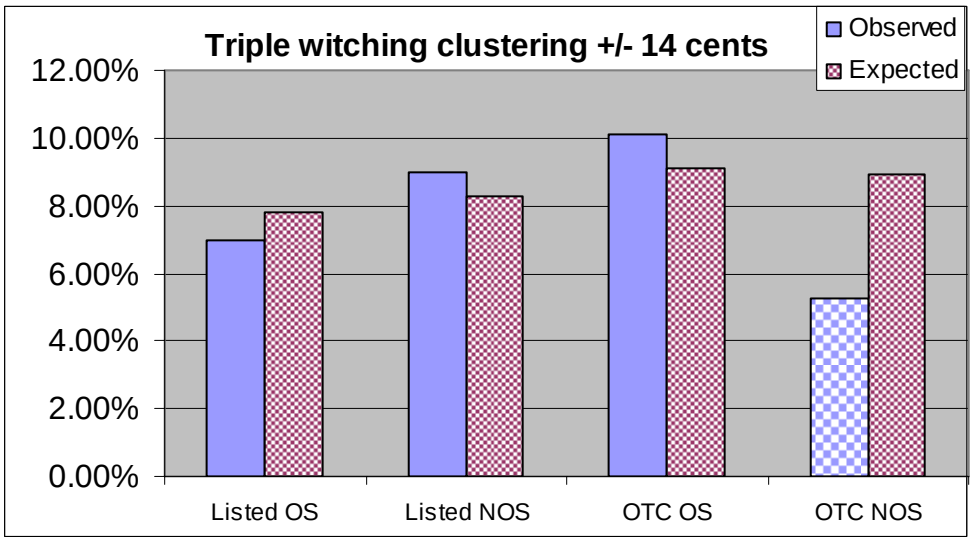


Figure 2B: Illustrates the triple witching clustering for listed OS and NOS, and OTC OS and NOS at the $[K \pm 14\,cents]$-interval. All OS were insignificantly different from the uniform distribution, whereas the OTC NOS experienced negative excess clustering (t = -2.26).

Appendix 3: Logistic regression

The results that were derived from the regression equations rely on a range of statistical assumptions. For example, it is assumed that the model is correctly specified. Failure of a "correctly" specified model leads to biased estimates (Johnston and DiNardo, 1998). If the models are misspecified, inconsistent probabilities between the Logit and the Probit can occur. Therefore, these probabilities were calculated and no significant difference appeared.

The standard errors of a Logit model are affected by heteroskedasticity, which is why robust standard errors were tried out post estimation. However, the use of robust standard errors had no impact on the rejection decisions, nor did any of the test-statistics become more or less significant.

Appendix 4: S&P 600: clustering across decimals

This section complements the discussion of Question 2 that reports on clustering on decimals and the 25-cent intervals. In this appendix, that discussion is extended to account for differences between OS and NOS on expiration days and non-expiration days. Test statistics and comments are provided in the captions of each figure.

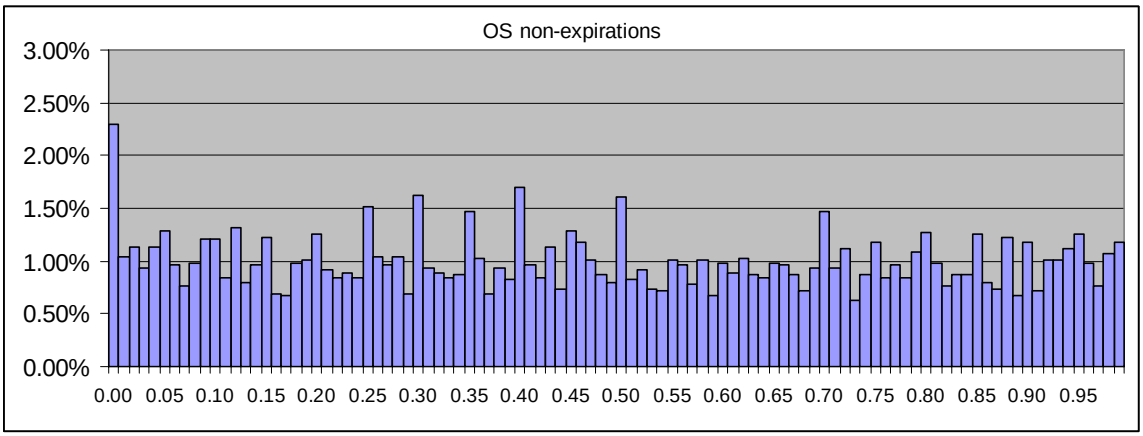


Figure 4A: Illustrates decimal-clustering for optionable stocks on Fridays before and Fridays after expiration Fridays. There were 4,486 stocks in this category and the Chi-square test statistic was 295.54, which rejects the uniform distribution at all levels.

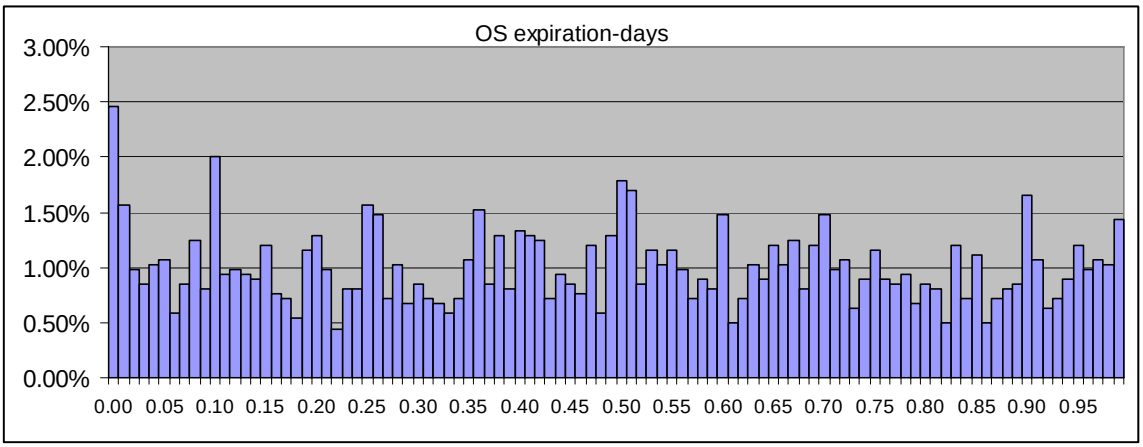


Figure 4B: Illustrates the clustering for optionable stocks on expiration Fridays. There were 2,241 observations in this category. The Chi-square test statistic against the uniform distribution was 256.14, whish also led to the rejection of the null hypothesis. The F-ratio between non-expiration and expiration clustering was 1.15, which was insufficient to differentiate between non-expiration and expiration day clustering.

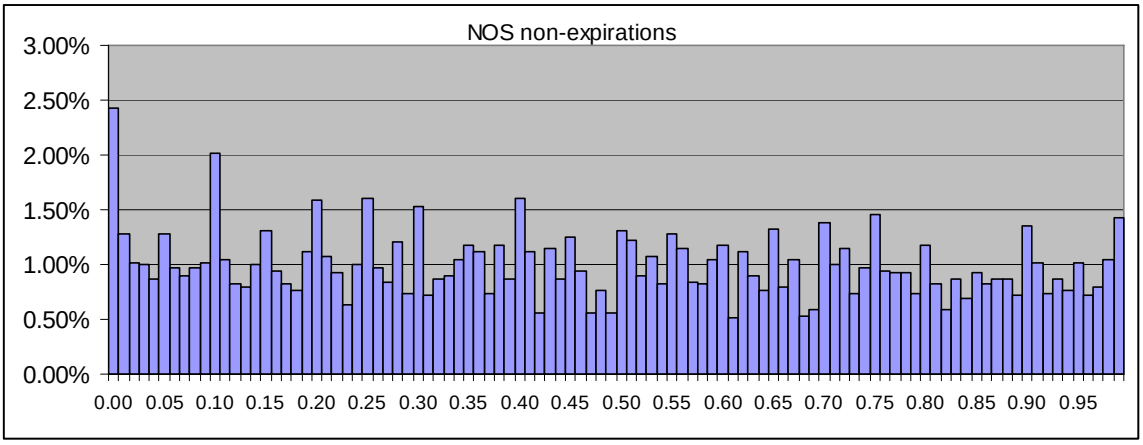


Figure 4C: Illustrates clustering for non-optionable stocks on Fridays before and Fridays after expiration Fridays. There were 3,914 stocks in this category and the Chi-square test statistic was 354.27, which implies rejection of the hypothesis that the distribution is uniform.

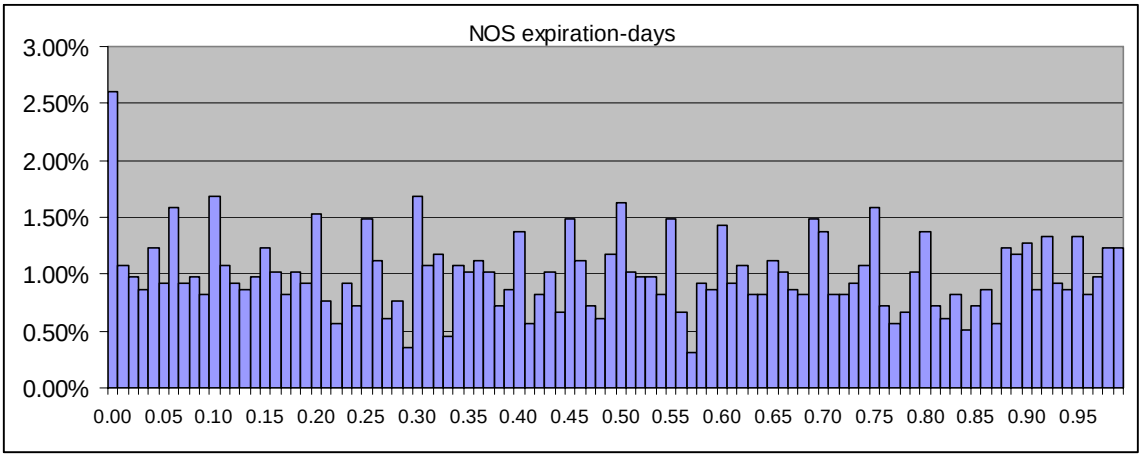


Figure 4D: Illustrates clustering for non-optionable stocks on expiration Fridays. There were 1,959 stocks in this category. The Chi-square tests statistic of the hypothesis that the decimal frequencies are uniformly distributed was 221.245. The F-ratio between non-expiration day and expiration day clustering was 1.60 (p-value = 0.01002), which rejects the hypothesis that the expiration-day and non-expiration day clustering for NOS is equal. Therefore, at the 5%-level, the expiration day clustering of NOS appears to be significantly more uniform than the non-expiration day clustering.

Appendix 5: An extension of the analysis in section 4.9

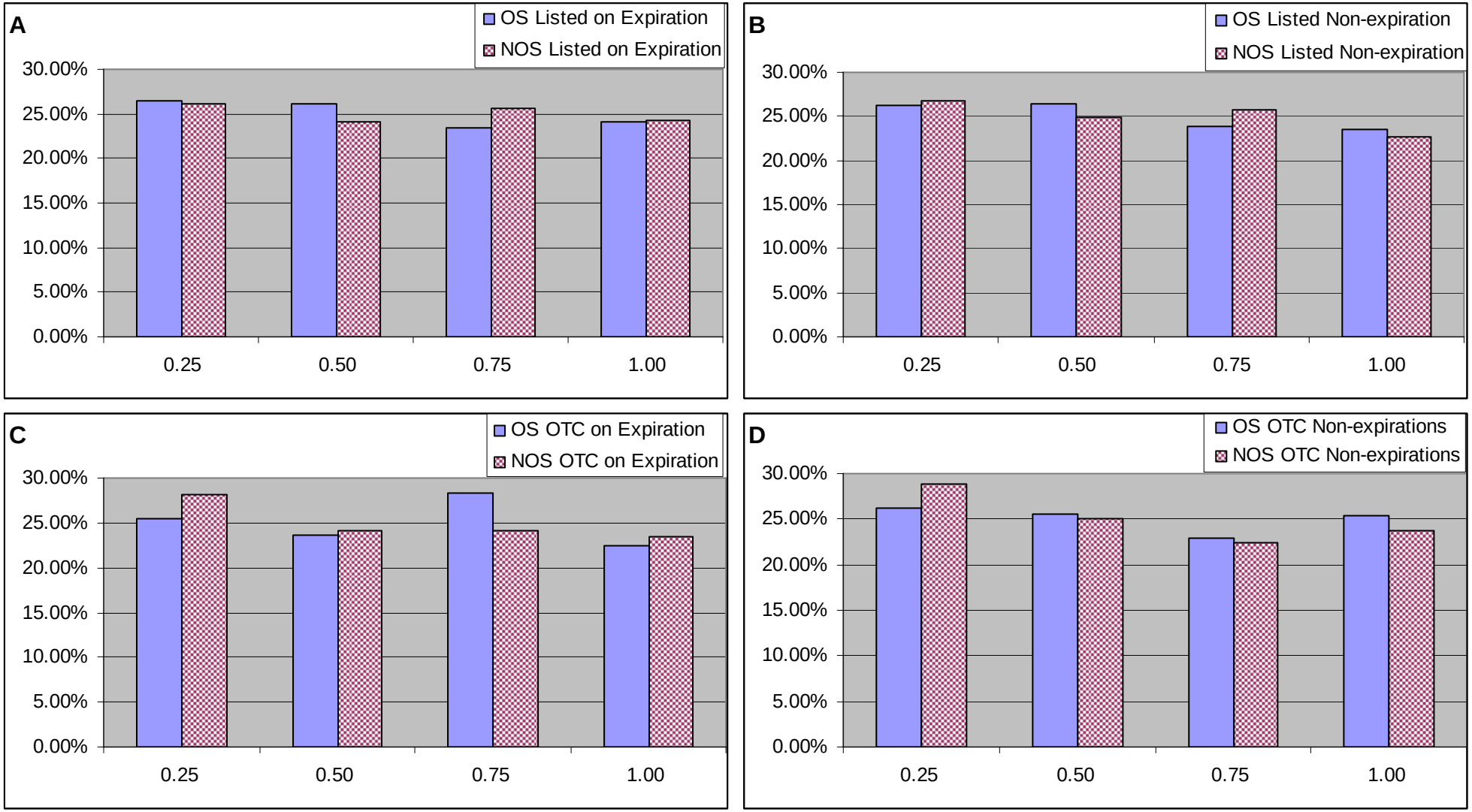


Figure 5A: An extension to the analysis in section 3.5 (NYSE, Nasdaq and AMEX clustering). In particular, panels A and B correspond to Figure 9 and panels C and D correspond to Figure 8. Only stocks in the S&P 600 index are included in this analysis. Panel A illustrates clustering for listed stocks on expiration days for OS and NOS. Both OS and NOS in panel A were insignificantly different from the uniform distribution in which the Chi-square test statistics were 2.52 and 1.50, respectively. Panel B represent the same stocks on non-expiration days. The Chi-square test statistics for OS and NOS were 5.04 and 8.96** (0.0299), respectively. Hence, on non-expiration days, NOS were non-uniform at the 5%-level. Panel C shows expiration day clustering for OTC stocks. The Chi-square test statistics for OS and NOS were 9.73** (p-value = 0.021) and 4.24, respectively. Hence, optionable OTC stocks were significantly different from the uniform distribution at expiration days. Panel D shows the clustering for the OTC stocks on non-expiration days. The Chi-square test statistics were 6.77* (p-value = 0.08) and 14.35*** (p-value = 0.0025), respectively. Hence, OS OTC stocks became uniform on non-expiration days, whereas NOS OTC stocks became non-uniform on non-expiration days. In general, however, listed stocks were distributed much more uniformly than OTC stocks; and when listed stocks became increasingly uniform on expiration days, OS OTC stocks became less uniform on expiration days.

Appendix 6: STATA CODE – Do-file

```
***A MASTER OF COMMERCE IN FINANCE

***An analysis of clustering for the S&P 100, 600 and Nasdaq 100
***August 2006 Oyvinn Rimer
***This file includes: Summary statistics, tests for clustering around
option strike
***prices and regression output from Logit and Probit regressions.
***Version 9.0
clear
set more off
set mem 200m


use "E:\Masters\Data\Analysis\By index\Pothsman Analysis\All new
stocks.dta", clear
log using "E:\Masters\Data\Analysis\By index\Pothsman Analysis\master
logfile.log", text replace


           *************
      ************************
   ******* SUMMARY STATISTICS *******
        ************************
           *************

*****************************************************************
*****************************************************************

********All indices********

***Optionable stocks w/volume: All indices***

preserve
drop if  dum_o==0|  dum_z_v==1|  dum_z_v==.
summ  expected_c_24 clustering_24 expected_c_14 clustering_14

/*calculates the Z-statistic for the difference between observed and
expected clustering
at the 25bps and 15bps intervals, respectively*/

**+/- 25
display (.1615464-.1541508)/(sqrt(.1541508*(1-.1541508)/10709))
ttest clustering_24 == expected_c_24, unpaired unequal
**+/- 15
display (0.1021571-.0912321)/(sqrt(.0912321*(1-.0912321)/10709))
ttest clustering_14 == expected_c_14, unpaired unequal

/*Hence, the clustering is significant at the larger interval, but not on
the
narrower interval*/
restore

***Optionable Friday before expiration***
preserve
drop if  dum_o==0|  dum_z_v==1|  dum_z_v==.
drop if dum_expi==1| after==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14
**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal
**+/- 15
```

```
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***Optionable expiration Friday ***
preserve
drop if  dum_o==0|  dum_z_v==1|  dum_z_v==.
drop if before==1| after==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14
**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal
**+/- 15
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***Optionable Friday after expiration***
preserve
drop if  dum_o==0|  dum_z_v==1|  dum_z_v==.
drop if dum_expi==1| before==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14
**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal
**+/- 15
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***Non-optionable stocks and optionable stocks w/no volume: All
indices****

preserve
drop if  dum_o==1 & dum_z_v==0
summ  expected_c_24 clustering_24 expected_c_14 clustering_14

*calculates the Z-statistic for the difference between observed and
expected clustering
*at the 25bps and 15bps intervals, respectively

**+/- 25
display (.1313609-.145645)/(sqrt(.145645*(1-.145645)/5915))
ttest clustering_24 == expected_c_24, unpaired unequal
**+/- 15
display (.0772612-.0861981)/(sqrt(.0861981*(1-.0861981)/5915))
ttest clustering_14 == expected_c_14, unpaired unequal

/*The clustering is significantly different from the theoretical
clustering
at both intervals, which indicates that optionable stocks are far more
efficient than non-optionable stocks*/
restore

***Non-optionable: Friday before expiration***
preserve
drop if  dum_o==1 & dum_z_v==0
drop if dum_expi==1| after==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14
**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal
**+/- 15
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***Non-optionable: expiration Friday***
preserve
drop if  dum_o==1 & dum_z_v==0
```

```
drop if before==1| after==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14
**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal
**+/- 15
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***Non-optionable: Friday after expiration ***
preserve
drop if  dum_o==1 & dum_z_v==0
drop if dum_expi==1| before==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14
**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal
**+/- 15
ttest clustering_14 == expected_c_14, unpaired unequal
restore

*****************************************************************
*****************************************************************

*********S&P 100**********

***Optionable stocks w/volume: S&P 100 index***

preserve
drop if  dum_o==0|  dum_z_v==1|  dum_z_v==.
drop if dum_n100==1| dum_sp600==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14

*calculates the Z-statistic for the difference between observed and
expected clustering
*at the 25bps and 15bps intervals, respectively

**+/- 25
display (.173731-.1715385)/(sqrt(.1715385*(1-.1715385)/1911))
ttest clustering_24 == expected_c_24, unpaired unequal

**+/- 15
display (.1051805-.1015228)/(sqrt(.1015228*(1-.1015228)/1911))
ttest clustering_14 == expected_c_14, unpaired unequal

/*Hence, the clustering is insignificant at both intervals*/
restore

***Optionable: Friday before expiration

preserve
drop if  dum_o==0|  dum_z_v==1|  dum_z_v==.
drop if dum_n100==1| dum_sp600==1| after==1| dum_expi==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14

**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal

**+/- 15
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***Optionable: expiration Friday

preserve
```

```
drop if  dum_o==0|  dum_z_v==1|  dum_z_v==.
drop if dum_n100==1| dum_sp600==1| after==1| before==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14

**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal

**+/- 15
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***Optionable: Friday after expiration

preserve
drop if  dum_o==0|  dum_z_v==1|  dum_z_v==.
drop if dum_n100==1| dum_sp600==1| dum_expi==1| before==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14

**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal

**+/- 15
ttest clustering_14 == expected_c_14, unpaired unequal
restore


***Non-optionable stocks and optionable stocks w/no volume: S&P100
index****

preserve
drop if  dum_o==1 & dum_z_v==0
drop if dum_n100==1| dum_sp600==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14

*calculates the Z-statistic for the difference between observed and
expected clustering
*at the 25bps and 15bps intervals, respectively

**+/- 25
display (.1428571-.168)/(sqrt(.168*(1-.168)/21))
ttest clustering_24 == expected_c_24, unpaired unequal

**+/- 15
display (.047619-.0994286)/(sqrt(.0994286*(1-.0994286)/21))
ttest clustering_14 == expected_c_14, unpaired unequal

/*Also for these zero-volume stocks is the clustering insignificant.
However, the
number of observations (21) is too low to make these results reliable.
These results suggest
that high-cap stocks, such as those in the S&P 100, experience clustering
similar to the
uniform distribution*/
restore


***Non-optionable: Friday before expiration

preserve
drop if  dum_o==1 & dum_z_v==0
drop if dum_n100==1| dum_sp600==1| after==1| dum_expi==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14
```

```
**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal

**+/- 15
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***Non-optionable: expiration Friday

preserve
drop if  dum_o==1 & dum_z_v==0
drop if dum_n100==1| dum_sp600==1| after==1| before==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14

**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal

**+/- 15
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***Non-optionable: Friday after expiration

preserve
drop if  dum_o==1 & dum_z_v==0
drop if dum_n100==1| dum_sp600==1| dum_expi==1| before==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14

**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal

**+/- 15
ttest clustering_14 == expected_c_14, unpaired unequal
restore

*****************************************************************
*****************************************************************

*********NASDAQ 100**********

***Optionable stocks w/volume: NASDAQ 100 index***

preserve
drop if  dum_o==0|  dum_z_v==1|  dum_z_v==.
drop if  dum_sp100==1| dum_sp600==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14

*calculates the Z-statistic for the difference between observed and
expected clustering
*at the 25bps and 15bps intervals, respectively

**+/- 25
display (.1801062-.176514)/(sqrt(.176514*(1-.176514)/2071))
ttest clustering_24 == expected_c_24, unpaired unequal

**+/- 15
display (.1197489-.1044674)/(sqrt(.1044674*(1-.1044674)/2071))
ttest clustering_14 == expected_c_14, unpaired unequal

/*Hence, the clustering is insignificant at the wider interval, but
statistically significant
at the narrower interval. This may suggest that Nasdaq (OTC) stocks are
more susceptible
```

```
to anomalous clustering. This is explored in the text above.*/
restore

***Optionable: Friday before expiration

preserve
drop if  dum_o==0|  dum_z_v==1|  dum_z_v==.
drop if  dum_sp100==1| dum_sp600==1| dum_expi==1| after==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14

**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal

**+/- 15
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***Optionable: expiration Friday

preserve
drop if  dum_o==0|  dum_z_v==1|  dum_z_v==.
drop if  dum_sp100==1| dum_sp600==1| before==1| after==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14

**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal

**+/- 15
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***Optionable: Friday after expiration

preserve
drop if  dum_o==0|  dum_z_v==1|  dum_z_v==.
drop if  dum_sp100==1| dum_sp600==1| before==1| dum_expi==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14

**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal

**+/- 15
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***Non-optionable stocks and optionable stocks w/no volume: NASDAQ 100
index***

preserve
drop if  dum_o==1 & dum_z_v==0
drop if  dum_sp100==1| dum_sp600==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14

*calculates the Z-statistic for the difference between observed and
expected clustering
*at the 25bps and 15bps intervals, respectively

**+/- 25
display (.1904762-.1866667)/(sqrt(.1866667*(1-.1866667)/21))
ttest clustering_24 == expected_c_24, unpaired unequal

**+/- 15
display (.1428571-.1104762)/(sqrt(.11047627*(1-.1104762)/21))
```

```
ttest clustering_14 == expected_c_14, unpaired unequal

/*Also for these zero-volume stocks is the clustering insignificant.
However, the
number of observations (21) is too low to make these results reliable*/
restore

***Non-optionable: Friday before expiration

preserve
drop if  dum_o==1 & dum_z_v==0
drop if  dum_sp100==1| dum_sp600==1| dum_expi==1| after==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14

**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal

**+/- 15
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***Non-optionable: expiration Friday

preserve
drop if  dum_o==1 & dum_z_v==0
drop if  dum_sp100==1| dum_sp600==1| before==1| after==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14

**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal

**+/- 15
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***Non-optionable: Friday after expiration

preserve
drop if  dum_o==1 & dum_z_v==0
drop if  dum_sp100==1| dum_sp600==1| before==1| dum_expi==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14

**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal

**+/- 15
ttest clustering_14 == expected_c_14, unpaired unequal
restore

*****************************************************************
*****************************************************************

*********S&P 600**********

***Optionable stocks w/volume: S&P 600 index***

preserve
drop if  dum_o==0|  dum_z_v==1|  dum_z_v==.
drop if  dum_sp100==1| dum_n100==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14

*calculates the Z-statistic for the difference between observed and
expected clustering
```

```
*at the 25bps and 15bps intervals, respectively

**+/- 25
display (.152371-.1423265)/(sqrt(.1423265*(1-.1423265)/6727))
ttest clustering_24 == expected_c_24, unpaired unequal

**+/- 15
display (0.0958823-.0842341)/(sqrt(.0842341*(1-.0842341)/6727))
ttest clustering_14 == expected_c_14, unpaired unequal

/*Hence, the clustering is significant at the wider interval, but
statistically insignificant
at the narrower interval.*/
restore

***LISTED OS***
preserve
drop if  dum_o==0|  dum_z_v==1|  dum_z_v==.
drop if  dum_sp100==1| dum_n100==1
drop if otc==1
ttest clustering_24 == expected_c_24, unpaired unequal
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***OTC OS***
preserve
drop if  dum_o==0|  dum_z_v==1|  dum_z_v==.
drop if  dum_sp100==1| dum_n100==1
drop if otc==0
ttest clustering_24 == expected_c_24, unpaired unequal
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***Non-optionable stocks and optionable stocks w/no volume: S&P 600
index***

preserve
drop if  dum_o==1 & dum_z_v==0
drop if  dum_sp100==1| dum_n100==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14

*calculates the Z-statistic for the difference between observed and
expected clustering
*at the 25bps and 15bps intervals, respectively

**+/- 25
display (.1311085-.1454184)/(sqrt(.1454184*(1-.1454184)/5873))
ttest clustering_24 == expected_c_24, unpaired unequal

**+/- 15
display (0.0771326-.086064)/(sqrt(.086064*(1-.086064)/5873))
ttest clustering_14 == expected_c_14, unpaired unequal

/*For both intervals, the clustering is significantly different from the
uniform distribution. There is also a negative sign on the test statistic,
which indicates that the NOS experience significantly less clustering than
the efficiency benchmark (uniform distribution).*/
restore

***LISTED NOS***
preserve
drop if  dum_o==1 & dum_z_v==0
drop if  dum_sp100==1| dum_n100==1
```

```
drop if otc==1
ttest clustering_24 == expected_c_24, unpaired unequal
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***OTC NOS***
preserve
drop if  dum_o==1 & dum_z_v==0
drop if  dum_sp100==1| dum_n100==1
drop if otc==0
ttest clustering_24 == expected_c_24, unpaired unequal
ttest clustering_14 == expected_c_14, unpaired unequal
restore

*******************************************************************
***S&P 600 OS***

***Optionable: Friday before expiration***
preserve
drop if  dum_o==0|  dum_z_v==1|  dum_z_v==.
drop if dum_expi==1| after==1| dum_sp100==1| dum_n100==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14
**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal
**+/- 15
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***Optionable: expiration Friday ***
preserve
drop if  dum_o==0|  dum_z_v==1|  dum_z_v==.
drop if before==1| after==1| dum_sp100==1| dum_n100==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14
**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal
**+/- 15
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***Optionable: Friday after expiration***
preserve
drop if  dum_o==0|  dum_z_v==1|  dum_z_v==.
drop if dum_expi==1| before==1| dum_sp100==1| dum_n100==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14
**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal
**+/- 15
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***Listed Optionable: Friday before expiration***
preserve
drop if  dum_o==0|  dum_z_v==1|  dum_z_v==.
drop if dum_expi==1| after==1| dum_sp100==1| dum_n100==1| otc==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14
**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal
**+/- 15
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***Listed Optionable: expiration Friday ***
preserve
```

```
drop if  dum_o==0|  dum_z_v==1|  dum_z_v==.
drop if before==1| after==1| dum_sp100==1| dum_n100==1| otc==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14
**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal
**+/- 15
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***Listed Optionable: Friday after expiration***
preserve
drop if  dum_o==0|  dum_z_v==1|  dum_z_v==.
drop if dum_expi==1| before==1| dum_sp100==1| dum_n100==1| otc==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14
**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal
**+/- 15
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***OTC Optionable: Friday before expiration***
preserve
drop if  dum_o==0|  dum_z_v==1|  dum_z_v==.
drop if dum_expi==1| after==1| dum_sp100==1| dum_n100==1| otc==0
summ  expected_c_24 clustering_24 expected_c_14 clustering_14
**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal
**+/- 15
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***OTC Optionable: expiration Friday ***
preserve
drop if  dum_o==0|  dum_z_v==1|  dum_z_v==.
drop if before==1| after==1| dum_sp100==1| dum_n100==1| otc==0
summ  expected_c_24 clustering_24 expected_c_14 clustering_14
**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal
**+/- 15
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***OTC Optionable: Friday after expiration***
preserve
drop if  dum_o==0|  dum_z_v==1|  dum_z_v==.
drop if dum_expi==1| before==1| dum_sp100==1| dum_n100==1| otc==0
summ  expected_c_24 clustering_24 expected_c_14 clustering_14
**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal
**+/- 15
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***S&P 600 NOS***

***Non-optionable: Friday before expiration***
preserve
drop if  dum_o==1 & dum_z_v==0
drop if dum_expi==1| after==1| dum_sp100==1| dum_n100==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14
**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal
**+/- 15
```

```
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***Non-optionable: expiration Friday***
preserve
drop if  dum_o==1 & dum_z_v==0
drop if before==1| after==1| dum_sp100==1| dum_n100==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14
**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal
**+/- 15
ttest clustering_14 == expected_c_14, unpaired unequal
*Z-test:
display (0.0755487-.0860436)/(sqrt(.0860436*(1-.0860436)/1959))
restore

***Non-optionable: Friday after expiration ***
preserve
drop if  dum_o==1 & dum_z_v==0
drop if dum_expi==1| before==1| dum_sp100==1| dum_n100==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14
**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal
**+/- 15
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***Listed Non-optionable: Friday before expiration***
preserve
drop if  dum_o==1 & dum_z_v==0
drop if dum_expi==1| after==1| dum_sp100==1| dum_n100==1| otc==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14
**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal
**+/- 15
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***Listed Non-optionable: expiration Friday***
preserve
drop if  dum_o==1 & dum_z_v==0
drop if before==1| after==1| dum_sp100==1| dum_n100==1| otc==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14
**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal
**+/- 15
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***Listed Non-optionable: Friday after expiration ***
preserve
drop if  dum_o==1 & dum_z_v==0
drop if dum_expi==1| before==1| dum_sp100==1| dum_n100==1| otc==1
summ  expected_c_24 clustering_24 expected_c_14 clustering_14
**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal
**+/- 15
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***OTC Non-optionable: Friday before expiration***
preserve
drop if  dum_o==1 & dum_z_v==0
```

```
drop if dum_expi==1| after==1| dum_sp100==1| dum_n100==1| otc==0
summ  expected_c_24 clustering_24 expected_c_14 clustering_14
**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal
**+/- 15
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***OTC Non-optionable: expiration Friday***
preserve
drop if  dum_o==1 & dum_z_v==0
drop if before==1| after==1| dum_sp100==1| dum_n100==1| otc==0
summ  expected_c_24 clustering_24 expected_c_14 clustering_14
**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal
**+/- 15
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***OTC Non-optionable: Friday after expiration ***
preserve
drop if  dum_o==1 & dum_z_v==0
drop if dum_expi==1| before==1| dum_sp100==1| dum_n100==1| otc==0
summ  expected_c_24 clustering_24 expected_c_14 clustering_14
**+/- 25
ttest clustering_24 == expected_c_24, unpaired unequal
**+/- 15
ttest clustering_14 == expected_c_14, unpaired unequal
restore

*******************************************************************
*******************************************************************
*** S&P 600 Triple Witching***

***LISTED OS***
preserve
use "E:\Masters\Data\Analysis\By index\Pothsman Analysis\Triple
witching.dta", clear
drop if  dum_o==0|  dum_z_v==1|  dum_z_v==.
drop if otc==1
ttest clustering_24 == expected_c_24, unpaired unequal
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***OTC OS***
preserve
use "E:\Masters\Data\Analysis\By index\Pothsman Analysis\Triple
witching.dta", clear
drop if  dum_o==0|  dum_z_v==1|  dum_z_v==.
drop if otc==0
ttest clustering_24 == expected_c_24, unpaired unequal
ttest clustering_14 == expected_c_14, unpaired unequal
restore

***LISTED NOS***
preserve
use "E:\Masters\Data\Analysis\By index\Pothsman Analysis\Triple
witching.dta", clear
drop if  dum_o==1 & dum_z_v==0
drop if otc==1
ttest clustering_24 == expected_c_24, unpaired unequal
ttest clustering_14 == expected_c_14, unpaired unequal
restore
```

```
***OTC NOS***
preserve
use "E:\Masters\Data\Analysis\By index\Pothsman Analysis\Triple
witching.dta", clear
drop if  dum_o==1 & dum_z_v==0
drop if otc==0
ttest clustering_24 == expected_c_24, unpaired unequal
ttest clustering_14 == expected_c_14, unpaired unequal
restore

*******************************************************
***************   5-CENT ANALYSIS   *******************
*******************************************************

*********GENERATING DUMMY FOR 5-CENT CLUSTERING********

*gen clustering_05 = abs(price-clustering_strike)<0.05
*gen expected_c_05=0.09/ delta

***test of clustering within 5-cents of an option strike price for all
stocks

ttest clustering_05 == expected_c_05, unpaired unequal

***ALL SP600 -05<K<05***
preserve
drop if dum_sp100==1| dum_n100==1
ttest clustering_05 == expected_c_05, unpaired unequal
restore

***Expiration day SP600***
preserve
drop if dum_sp100==1| dum_n100==1
drop if before==1| after==1
ttest clustering_05 == expected_c_05, unpaired unequal
restore

***Expiration day SP600 LISTED OPTIONABLE STOCKS***
preserve
drop if dum_sp100==1| dum_n100==1
drop if before==1| after==1
drop if  dum_o==0|  dum_z_v==1|  dum_z_v==.
drop if otc==1
ttest clustering_05 == expected_c_05, unpaired unequal
restore

***Expiration day SP600 LISTED NON-OPTIONABLE STOCKS***
preserve
drop if dum_sp100==1| dum_n100==1
drop if before==1| after==1
drop if  dum_o==1 & dum_z_v==0
drop if otc==1
ttest clustering_05 == expected_c_05, unpaired unequal
restore

***Expiration day SP600 OTC OPTIONABLE STOCKS***
preserve
drop if dum_sp100==1| dum_n100==1
drop if before==1| after==1
drop if  dum_o==0|  dum_z_v==1|  dum_z_v==.
drop if otc==0
ttest clustering_05 == expected_c_05, unpaired unequal
```

```
restore

***Expiration day SP600 OTC NON-OPTIONABLE STOCKS***
preserve
drop if dum_sp100==1| dum_n100==1
drop if before==1| after==1
drop if  dum_o==1 & dum_z_v==0
drop if otc==0
ttest clustering_05 == expected_c_05, unpaired unequal
restore

***Before expiration S&P600***
preserve
drop if dum_sp100==1| dum_n100==1
drop if dum_expi==1| after==1
ttest clustering_05 == expected_c_05, unpaired unequal
restore

***Before expiration day SP600 LISTED OPTIONABLE STOCKS***
preserve
drop if dum_sp100==1| dum_n100==1
drop if dum_expi==1| after==1
drop if  dum_o==0|  dum_z_v==1|  dum_z_v==.
drop if otc==1
ttest clustering_05 == expected_c_05, unpaired unequal
restore

***Before expiration day SP600 LISTED NON-OPTIONABLE STOCKS***
preserve
drop if dum_sp100==1| dum_n100==1
drop if dum_expi==1| after==1
drop if  dum_o==1 & dum_z_v==0
drop if otc==1
ttest clustering_05 == expected_c_05, unpaired unequal
restore

***Before expiration day SP600 OTC OPTIONABLE STOCKS***
preserve
drop if dum_sp100==1| dum_n100==1
drop if dum_expi==1| after==1
drop if  dum_o==0|  dum_z_v==1|  dum_z_v==.
drop if otc==0
ttest clustering_05 == expected_c_05, unpaired unequal
restore

***Before expiration day SP600 OTC NON-OPTIONABLE STOCKS***
preserve
drop if dum_sp100==1| dum_n100==1
drop if dum_expi==1| after==1
drop if  dum_o==1 & dum_z_v==0
drop if otc==0
ttest clustering_05 == expected_c_05, unpaired unequal
restore

***After expiration S&P600***
preserve
drop if dum_sp100==1| dum_n100==1
drop if dum_expi==1| before==1
ttest clustering_05 == expected_c_05, unpaired unequal
restore

***AFTER expiration day SP600 LISTED OPTIONABLE STOCKS***
preserve
```

```
drop if dum_sp100==1| dum_n100==1
drop if dum_expi==1| before==1
drop if  dum_o==0|  dum_z_v==1|  dum_z_v==.
drop if otc==1
ttest clustering_05 == expected_c_05, unpaired unequal
restore

***AFTER expiration day SP600 LISTED NON-OPTIONABLE STOCKS***
preserve
drop if dum_sp100==1| dum_n100==1
drop if dum_expi==1| before==1
drop if  dum_o==1 & dum_z_v==0
drop if otc==1
ttest clustering_05 == expected_c_05, unpaired unequal
restore

***AFTER expiration day SP600 OTC OPTIONABLE STOCKS***
preserve
drop if dum_sp100==1| dum_n100==1
drop if dum_expi==1| before==1
drop if  dum_o==0|  dum_z_v==1|  dum_z_v==.
drop if otc==0
ttest clustering_05 == expected_c_05, unpaired unequal
restore

***AFTER expiration day SP600 OTC NON-OPTIONABLE STOCKS***
preserve
drop if dum_sp100==1| dum_n100==1
drop if dum_expi==1| before==1
drop if  dum_o==1 & dum_z_v==0
drop if otc==0
ttest clustering_05 == expected_c_05, unpaired unequal
restore

***ALL NASDAQ100 -05<K<05***
preserve
drop if dum_sp100==1| dum_sp600==1
ttest clustering_05 == expected_c_05, unpaired unequal
restore

***Expiration day N100***
preserve
drop if dum_sp100==1| dum_sp600==1
drop if before==1| after==1
ttest clustering_05 == expected_c_05, unpaired unequal
restore

***Before expiration N100***
preserve
drop if dum_sp100==1| dum_sp600==1
drop if dum_expi==1| after==1
ttest clustering_05 == expected_c_05, unpaired unequal
restore

***After expiration N100***
preserve
drop if dum_sp100==1| dum_sp600==1
drop if dum_expi==1| before==1
ttest clustering_05 == expected_c_05, unpaired unequal
restore

***ALL SP100 -05<K<05***
preserve
```

```
drop if dum_n100==1| dum_sp600==1
ttest clustering_05 == expected_c_05, unpaired unequal
restore

***Expiration day SP100***
preserve
drop if dum_n100==1| dum_sp600==1
drop if before==1| after==1
ttest clustering_05 == expected_c_05, unpaired unequal
restore

***Before expiration SP100***
preserve
drop if dum_n100==1| dum_sp600==1
drop if dum_expi==1| after==1
ttest clustering_05 == expected_c_05, unpaired unequal
restore

***After expiration SP100***
preserve
drop if dum_n100==1| dum_sp600==1
drop if dum_expi==1| before==1
ttest clustering_05 == expected_c_05, unpaired unequal
restore

*****************************************************************************
*****************************************************************************
*****************************************************************************

***Logistic regression of clustering***

***Generates logged variables***
gen ln_call_vol=log(call_vol)
gen ln_call_oi=log(call_oi)
gen ln_put_vol=log(put_vol)
gen ln_put_oi=log(put_oi)
gen ln_price=log(price)

preserve
drop if dum_sp100==1| dum_n100==1

logit clustering_24 ln_call_vol ln_call_oi ln_put_vol ln_put_oi ln_price
otc, robust
mfx compute

logit clustering_14 ln_call_vol ln_call_oi ln_put_vol ln_put_oi ln_price
otc, robust
lroc

*mfx compute

/*reg clustering_14 ln_call_vol ln_call_oi ln_put_vol ln_put_oi ln_price
otc
mfx compute

logit clustering_05 ln_call_vol ln_call_oi ln_put_vol ln_put_oi ln_price
otc
mfx compute


***OLS REGRESSION OF LOG(S/K)***
gen Abs_ln_S_over_K=abs(log(price/clustering_strike))
gen ln_S_over_K=log(price/clustering_strike)
```

```
reg Abs_ln_S_over_K ln_call_vol ln_call_oi ln_put_vol ln_put_oi ln_price
otc
reg ln_S_over_K ln_call_vol ln_call_oi ln_put_vol ln_put_oi ln_price otc
restore*/

/*********************************************************************
*********************************************************************
*********************************************************************

***Ikenberry Analysis***
***Generates a series that classifies all stocks according to their
decimal bins.

gen round=int(price)
gen decimals=round(price-round,.01)
*hist decimals, bin(100)


***SP600 LISTED OPTIONABLE STOCKS***
preserve
drop if dum_sp100==1| dum_n100==1
drop if  dum_o==0|  dum_z_v==1|  dum_z_v==.
drop if otc==1
*hist decimals, bin(100)
restore

***SP600 LISTED NON-OPTIONABLE STOCKS***
preserve
drop if dum_sp100==1| dum_n100==1
drop if  dum_o==1 & dum_z_v==0
drop if otc==1
*hist decimals, bin(100)
restore

***SP600 OTC OPTIONABLE STOCKS***
preserve
drop if dum_sp100==1| dum_n100==1
drop if  dum_o==0|  dum_z_v==1|  dum_z_v==.
drop if otc==0
*hist decimals, bin(100)
restore

***SP600 OTC NON-OPTIONABLE STOCKS***
preserve
drop if dum_sp100==1| dum_n100==1
drop if  dum_o==1 & dum_z_v==0
drop if otc==0
*hist decimals, bin(100)
restore
*/
log close
```