University of Alberta

Three Essays in Finance

by

Kenton Karl Hoyem ©

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Finance

Faculty of Business

Edmonton, Alberta
Fall 2008

Your file Votre référence
ISBN: 978-0-494-46336-9
Our file Notre référence
ISBN: 978-0-494-46336-9

University of Alberta

Library Release Form

Name of Author: Kenton Karl Hoyem

Title of Thesis: Three Essays in Finance

Degree: Doctor of Philosophy

Year this Degree Granted: 2008

Signature

Any sufficiently advanced technology is indistinguishable from magic.

- *Arthur C. Clark*

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **Three Essays in Finance** submitted by **Kenton Karl Hoyem** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy**.

Jung-Wook Kim, Ph.D.

Randall Morck, Ph.D.

Vikas Mehrotra, Ph.D.

Constance Smith, Ph.D.

Alfred Lehar, Ph.D.

DEDICATION

To I.H. and S.J.

ABSTRACT

This thesis presents three essays on trading volume, capital gains overhang (CGO) and analysts' recommendations and estimates. In the first essay, I examine a possible cause for the higher returns experienced by stocks with higher volume around earnings announcements. I find that these returns are concentrated in stocks with either large aggregate unrealized capital gains or losses. A high volume minus low volume portfolio conditioned on the magnitude of CGO generates significant and robust returns as high as 11% per year. This suggests that this earnings announcement volume premium is associated with selling pressure from investors who are influenced by the magnitude of unrealized capital gains or losses. This also suggests that the well known disposition effect may not hold for stocks with extreme unrealized capital losses and is consistent with recent theoretical and empirical research that shows extreme losses prompt selling.

In the second essay, I analyze the accuracy of the modified Grinblatt and Han (2005) CGO approximation methodology (employed in the first essay) by utilizing a transaction-level data set. I find that their first simplifying assumption, using daily closing prices instead of actual transaction prices, increases daily CGO values by approximately 0.11%, but that this increase occurs uniformly across observations and does not significantly affect the relative ranking of CGO values. I also find that their second simplifying assumption, using a weighted average method of calculating CGO values, is superior to both first-in-first-out and last-in-first-out methods.

In the third essay, I examine analysts' recommendations and estimates. Analysts are generally viewed as having a positive or optimistic bias in their stock recommendations. Additionally, it is widely believed that analysts will drop coverage of firms instead of issuing highly negative recommendations or estimates. I evaluate whether the bias of analysts' consensus recommendations and annual earnings estimates can be corrected through adjustments for dropped coverage, and if such corrections increase recommendation and estimate accuracy. I find that corrective adjustments can significantly reduce, although not eliminate, the optimism bias for recommendations. Such corrections do not significantly improve the accuracy of recommendations, but they can significantly improve the accuracy of annual earnings estimates.

ACKNOWLEDGEMENTS

I wish to express my sincere thanks and appreciation to my supervisor Professor Jung-Wook Kim. This thesis would not have been possible without his wise guidance and infinite patience.

I also wish to thank Dr. Wonseok Choi, Professors Randall Morck, Vikas Mehrotra, Constance Smith, Nadia Massoud and Alfred Lehar for their efforts in seeing my thesis come to completion. Additionally, I wish to acknowledge the other faculty members and my fellow graduate students in the Ph.D. program at the University of Alberta, as my work has greatly benefited from their numerous comments and suggestions.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

ACP:	Actual closing price
ATP:	Actual transaction price
CB:	Cost basis
CBA:	Cost basis adjustment
CGO:	Capital gains overhang
CPS:	Cost per share
DRIFT60:	Sum of daily size- and B/M-adjusted returns from days $t+2$ to $t+61$
EAVP:	Earnings announcement volume premium
EPS:	Earnings per share
ESTTO:	Estimated daily turnover
FIFO:	First-in-First-Out
LIFO:	Last-in-First-Out
LSIZE_SUE:	SUE interacted with the natural logarithm of company size
m:	Time (month)
MKTTO:	Value weighted turnover for the entire market
PEAD:	Post earnings announcement drift
SB:	Share balance
SBA:	Share balance adjustment
SUE:	Standardized unexpected earnings
t:	Time (day)
TO:	Daily turnover
VOL:	Abnormal volume decile number
VOL_CGO:	VOL interacted with CGO
VOL_CGONEG:	Absolute value of negative CGO interacted with VOL
VOL_CGOPOS:	Absolute value of positive CGO interacted with VOL
WA:	Weighted average
y:	Time (year)

CHAPTER 1

INTRODUCTION

In this thesis I present three essays on topics in finance. In Chapter 2, I examine the high volume return premium around earnings announcements and its interaction with capital gains overhang (i.e. unrealized capital gains or losses). In Chapter 3, I further examine the capital gains approximation methodology used in Chapter 2, and evaluate its accuracy versus alternate methodologies. In Chapter 4, I examine the impact of dropped coverage on the accuracy of analysts' recommendations and earnings estimates, and evaluate modification methods for improving the accuracy of both recommendations and estimates which take into account dropped coverage.

Chapter 2 examines the high volume return premium in relation to capital gains overhang. The high volume return premium is a well-documented phenomenon where stocks which have high abnormal volume over a given period tend to exhibit higher subsequent abnormal returns. I hypothesize that this is attributable to the buying and selling of stocks which have either large capital gains or losses, and that the actions of sellers is driven by behavioural biases. Specifically, investors will become increasingly risk averse when sitting on large unrealized capital gains, and will make potentially irrational selling decisions which results in a temporary downward pressure on the stock price. This will be followed by subsequent high returns as the stock adjusts to its 'rational' price. In the case of large capital losses, investors who were previously reluctant to consummate their losses will eventually be forced to realize them after a given point, whether due to capital constraints or other factors. This will also result in a temporary downward pressure on the stock price, followed by a subsequent rebound.

To test this hypothesis, I examine the high volume return premium around quarterly earnings announcements, or the earnings announcement volume premium. Earnings announcements are significant news events, and lead to an increase in trading volume – as such, they provide a fertile ground for conducting a volume related study. Consistent with prior research, I find the presence of an earnings announcement volume premium.

I then calculate aggregate capital gains overhang measures for individual stocks using a modified version of Grinblatt and Han's (2005) methodology. This allows for the estimation of the aggregate (i.e. market-wide) capital gains overhang for individual stocks. The earnings announcement volume premium results are further conditioned on this capital gains overhang value, and I find that the earnings announcement volume premium is primarily concentrated in stocks with either high capital gains or losses, and is insignificant for those with low or negligible capital gains or losses. I interpret this as evidence in support of my hypothesis that behavioural factors play a significant role in explaining the earnings announcement volume premium.

Chapter 3 evaluates the accuracy of the modified Grinblatt and Han (2005) capital gains overhang methodology used in Chapter 2. One obstacle faced by researchers focusing on capital gains is the lack of transaction-level data at the individual investor level. As a result, methodologies have been developed to approximate the aggregate level of capital gains overhang using daily closing prices and turnover values. In the Grinblatt and Han (2005) methodology, it is assumed that all existing investors sell to new investors in a pro-rata or weighted average fashion, and that all transactions occur at the day's closing price.

Using a transaction-level data set from an American brokerage firm, I test these assumptions by calculating actual capital gains overhang values at the level of individual investors for each stock. These are then aggregated and the results compared to values calculated using the modified Grinblatt and Han (2005) methodology on the same data set. Additionally, I evaluate whether investors' behaviour is better modelled using a *first-in-first-out* methodology (where initial investors are assumed to sell their holdings first) or a *last-in-first-out* methodology (where the most recent investors are assumed to sell their holdings first).

I find that using daily closing prices instead of actual transaction prices slightly increases capital gains overhang values, but that this increase is uniform across the distribution of capital gains overhang values, and as such does not introduce major distortions to any tercile or quintile categorizations. Additionally, I find that the weighted average methodology, while not perfect, does provide a superior approximation versus either the *first-in-first-out* or the *last-in-first-out* methodologies. Overall, I conclude that the modified Grinblatt and Han (2005) methodology for estimating aggregate capital gains overhang provides reasonable approximations, particularly for calculating the relative capital gains overhang values of individual stocks.

Chapter 4 examines analysts' consensus recommendations and earnings estimates. Consistent with prior research, I find that there is a positive bias in the recommendations. This positive bias has typically been ascribed to overly optimistic recommendations by analysts, which could be due to economic factors such as existing business relationships with the covered company, or behavioural factors such as overreaction to good news or analyst herding. While not discounting these explanations, I seek to supplement the existing research by considering not only those analysts who are issuing recommendations or estimates, but also those analysts who have stopped issuing recommendations or estimates. I hypothesize that pessimistic analysts are removing themselves from the analyst pool rather than issuing a highly negative recommendation or estimate, and as a result the remaining recommendations and estimates demonstrate a positive bias.

I empirically test this hypothesis using both consensus recommendations and individual analysts' annual earnings per share estimates. In both cases, I use a variety of methods to adjust for dropped coverage, all of which use a pessimistic recommendation or estimate to replace dropped values. In the case of consensuses recommendations I find that the bias can be reduced through such methods, but that the accuracy is not significantly improved. For annual earnings per share estimates, I find that the accuracy can be significantly improved, with a rough approximation being that a dropping analyst can be considered to be issuing an estimate equal to the minimum of the remaining estimates for the following three years.

BIBLIOGRAPHY

Grinblatt, Mark, and Bing Han, 2005, Prospect Theory, Mental Accounting, and Momentum, *Journal of Financial Economics* 78, 311-339.

CHAPTER 2

CAPITAL GAINS OVERHANG AND THE EARNINGS ANNOUNCEMENT VOLUME PREMIUM[1]

2.1 INTRODUCTION

Earnings announcements lead to a substantial increase in trading volume (Lee *et al.* (1993), Kandel and Pearson (1995)), and there is a positive relationship between this trading volume and both contemporaneous and future stock returns (Bamber and Cheon (1995), Gervais *et al.* (2001), Garfinkel and Sokobin (2006), Lamont and Frazzini (2007), and Lerman *et al.* (2008)). In this paper, we present a novel hypothesis that implies the positive relationship between trading volume and future returns will be stronger in subsets of stocks with either large capital gains or losses. We report robust empirical findings that support the hypothesis.

Lamont and Frazzini (2007) report that stocks whose trading activity is mainly concentrated around earnings announcements attract higher returns in months where earnings announcements are expected. Their finding is an extension of the contemporaneous positive relationship between trading volume and stock returns reported in Karpoff (1987) and Bamber and Cheon (1995).[2] Lamont and Frazzini (2007) ascribe this to temporary buying pressure associated with increased investor attention (Barber and Odean (2008)) in the expected month of earnings announcement.

Other studies suggest that price increases after volume shocks may persist for a sustained period of time. Gervais *et al.* (2001) find that stocks experiencing positive

[1] This chapter is based upon the paper "Not All Trading Volumes are Created Equal: Capital Gains Overhang and the Earnings Announcement Volume Premium" co-authored with Wonseok Choi and Jung-Wook Kim, which has been submitted for publication.

[2] Bamber and Cheon (1995) report that firms with large trading volume and small price changes around earnings announcements are more likely to be associated with price increases than are firms with small trading volume and large price changes.

volume shocks tend to outperform those with negative shocks over the next 50 to 100 trading days.[3] This high volume-high return relationship is also found in a wide range of countries (Kaniel *et al.* (2005)).[4] Garfinkel and Sokobin (2006) and Lerman *et al.* (2008) also find that stocks with high abnormal volume measured around earnings announcements experience stronger drift (as measured over the following 60 trading days or until one day after the subsequent earnings announcement, respectively).

Trading requires both buyers and sellers. Thus, to explain the relationship between trading volume and returns, we need to explain the motivations of both groups. Lamont and Frazzini (2007) suggest that a short lived (within the expected month of earnings announcement) effect of trading activity around earnings announcements may reflect (possibly irrational) attention driven buying, which is accommodated by liquidity providers or arbitrageurs who exploit temporary overpricing. However, a price increase accompanied by heavy trading may not reverse quickly, or increased trading may be associated with continued price increases over a relatively long horizon, as in Gervais *et al.* (2001), Garfinkel and Sokobin (2006), and Lerman *et al.* (2008). In such cases, we require an additional explanation regarding who initially sells stocks, because such sellers seem to lose money on average.

In this paper, we report robust empirical findings for a calendar time trading strategy based on abnormal volume measured in a three day window around earnings announcements, and present a possible explanation. We focus on abnormal trading

---

3 They argue that the positive relationship between volume and future returns is distinct from trading volume's impact on the autocorrelation pattern of stock returns (Campbell *et al.* (1993), Wang (1994), and Llorente *et al.* (2002)) by showing that even when large abnormal trading volume is observed with small price changes, stock prices continue to go up.

4 Most models of trading volume cannot explain how high abnormal volume or volume shock leads to *persistent* increases in prices. For example, Wang (1994) argues that if private information drives high volume, the impact of this volume on future prices depends on the price change that accompanies the high volume. In models where trading volume is driven by portfolio rebalancing activities (Lo and Wang (2000, 2006)), high volume does not necessarily predict higher future returns. Models where trading volume results from differences in opinions (Kandel and Pearson (1995)) cannot explain the positive relationship between current trading volume and future returns, either. On the contrary, there are several papers that suggest increases in prices with high trading volume will eventually revert. Miller (1977) conjectures that high dispersion in opinions will lead to lower returns, because the views of pessimists are not properly reflected in stock prices due to short sale constraints. Lamont and Frazzini (2007) suggest that if high dispersion in opinions generates high trading volume, then this volume can be associated with temporary overpricing of stocks. Baker and Stein (2004) argue that high trading volume indicates the presence of optimistic noise traders and show that such volume leads to lower returns.

volume around earnings announcements as many investors make substantial portfolio rebalancing decisions around earnings announcements.

Our strategy is to establish long (short) positions in stocks with high (low) abnormal volume from the first day of the next month after an earnings announcement and hold them until the end of the next earnings announcement month or for four months, whichever comes first. By construction, the mean (median) lag before being included in a portfolio is 14.42 (13.00) days after an earnings announcement. This strategy is different from Lamont and Frazzini (2007), where stocks are purchased on the first day of the expected month of earnings announcement and held for a month. Our strategy is based on actual earnings announcements, stocks are included with significant lags, and stocks are held for a longer period. We do this to examine the persistency of the impact of trading volume on future returns (i.e. those returns which are not confined to the immediate window around earnings announcements). Surprisingly, despite this lag, this strategy generates positive and significant profits ranging from approximately 5% to over 11% per year and survives a series of risk adjustments. This implies that the profits are not merely a reflection of a positive relationship between price and volume just around earnings announcements. We define the return of the zero investment portfolio as the earnings announcement volume premium (henceforth EAVP).

In identifying the source of the EAVP, we first examine whether known risk factors may explain the profit of the strategy. We examine risk adjusted profit based on Fama and French's 3-factor model and its extensions which include price momentum (Carhart (1997)) or liquidity (Sadka (2006)) factors.[5] To check whether these profits merely reflect the well known post-earnings announcement drift (Bernard and Thomas (1989, 1990), henceforth PEAD), we also include a SUE factor.[6] Our results are also robust to

---

[5] Chordia *et al.* (2007a) report that the increase in trading volume around earnings announcements is insignificant for the most illiquid stocks. Since we examine stocks with large abnormal volumes, it is unlikely that the returns we document are explained by the so called illiquidity premium (Amihud (2002)). Nevertheless, we do control for illiquidity in the following analyses.

[6] SUE is standardized unexpected earnings based on a seasonal random walk model. The SUE factor is calculated as the difference in monthly equally weighted returns between the highest and the lowest SUE decile portfolios. Full details of our calculation methodology can be found in Section 2.2.

this additional adjustment.[7] This is consistent with Lerman *et al.* (2008) who show that the high volume-high return relationship around earnings announcements and PEAD are not the same phenomenon.[8]

To explain the EAVP, we develop a testable hypothesis which focuses on the role of unrealized capital gains or losses. Our hypothesis is motivated by the well known disposition effect (Shefrin and Statman (1985)) and shows that the interaction between investors who base their selling decision on unrealized capital gains or losses and liquidity providers or arbitrageurs creates trading volume that has future return implications.

The disposition effect refers to certain investors' tendency to sell winners prematurely (due to increased risk aversion for gains) and to hold onto losers far too long (due to increased risk taking for losses). Premature selling pressure on the winner stocks would be accommodated by arbitrageurs (thus creating high trading volume) and the arbitrageurs would be rewarded with high abnormal returns afterwards. Lack of selling pressure for loser stocks would be associated with low trading volume and subsequent low abnormal returns. In a similar spirit, Grinblatt and Han (2005) suggest that the price-momentum effect (Jegadeesh and Titman (1993)) arises because of this disposition effect: stocks that are sold by investors realizing gains tend to be underpriced due to excessive selling pressure. Conversely, stocks for which investors are reluctant to realize losses tend to be overpriced due to reduced selling pressure. They find supporting evidence in the cross-section of stock returns. Using mutual fund holdings data, Frazzini (2006) also finds evidence that PEAD is most pronounced when capital gains and news events have the same sign. However, most of the studies examining the disposition effect, including

---

7 Even though the EAVP survives adjustments for known risk factors, it may still represent unknown risk factors. For example, Garfinkel and Sokobin (2006) suggest that trading volume may proxy for differences in opinion which could be a priced risk factor, which depresses stock prices, generating higher expected returns. However, other proxies for opinion divergence do not fare well. For example, Garfinkel and Sokobin (2006) report that high values of analysts' forecast dispersion do not lead to stronger earnings announcement drift, contrary to the risk based interpretation. Diether *et al.* (2002) also find a negative relationship between analysts' forecast dispersion and stock returns. Avramov *et al.* (2008) show that the negative effect of analysts' forecast dispersion is mainly due to low credit rated firms' subsequent price drops which accompanies increases in forecast dispersion. In any case, opinion divergence proxied by analysts' forecast dispersion is associated with decreases in prices rather than increases in prices, and thus will not be able to explain the EAVP.

8 For example, our result suggests that future returns of good news firms with large abnormal trading volume are different from those with small abnormal trading volume.

Frazzini (2006) who examines the effect around earnings announcements, focus on the relationship between past capital gains (or losses) and returns without investigating the role of trading volume explicitly, even though the disposition effect is supposed to influence prices only through abnormal trading decisions.

On the other hand, several recent theoretical and empirical works suggest that the reluctance to realize losses implied by the disposition effect may not hold for extreme losses. Unlike stocks that experience capital gains, wealth can be a binding constraint in the case of extreme capital losses. Simply put, investors may not accumulate losses indefinitely. Barberis and Huang (2001), Barberis *et al.* (2001), and Gomes (2005) present theoretical possibilities that investors become more risk averse after a continued drop in a stock's price and may eventually decide to realize the losses. Consistent with this, Teo and O'Connell (2008) find that institutional investors aggressively reduce risk following losses in currency trading. Chordia *et al.* (2007b) find that greater magnitudes of past returns, even when they are negative, increase trading activity, and argue that such patterns are due to the portfolio rebalancing needs of investors who are reacting to changes in asset valuations. Tax considerations are also known to induce loss realization. As losses deepen, the tax benefits of realizing losses may outweigh the tendency to hold onto losing stocks. Grinblatt and Keloharju (2001, 2004), Grinblatt and Moskowitz (2004), and Starks *et al.* (2006) provide evidence of tax-loss selling, which is more pronounced near the end of the year.[9] Jin (2006) also finds that capital gains taxes encourage institutions that serve tax sensitive clients to sell stocks with large capital losses around earnings announcements. All this evidence allows us to hypothesize that abnormal selling pressure may arise for stocks with large capital losses.

Among the possible reasons for selling stocks with capital losses, some could be consistent with rational investor behavior. For example, realizing losses for tax benefits or selling losing stocks to stop further losses. On the contrary, other reasons could be consistent with sub-optimal investor behavior. For example, investors may postpone

[9] Grinblatt and Moskowitz (2004) find that being a consistent winner has a substantial positive impact on the cross-section of returns (possibly due to premature realization of gains), while being a consistent loser appears to be irrelevant to the cross-section of returns. They suspect that the weak return predictability for consistent losers could be related to a tax-loss selling induced reversal in returns that offsets momentum in these stocks.

selling losing stocks far too long due to a reluctance to realize losses and eventually sell them at the worst moment (i.e. when the stocks have potential for price increases). Whatever the case may be, if selling pressures are not driven by forecasts of future fundamentals but driven by the magnitude of past losses only, liquidity providers or rational arbitrageurs may take the other side of such trades and the prices of those stocks will subsequently rebound.

Discussions so far suggest a testable hypothesis that the EAVP could arise as a result of interactions between two groups of investors: those who make their portfolio decisions with regard to capital gains or losses, and arbitrageurs who trade only if there is a profit opportunity. If unrealized capital gains or losses do not affect the trading decisions of investors, we do not expect to find any systematic relationship between the EAVP and capital gains overhang. However, if there are investors who are affected by capital gains overhang (due to the disposition effect or loss realization), the EAVP should be stronger in either or both of the two extremes of capital gains overhang.

To examine this issue, we divide our sample into quintiles based on the aggregate measure for unrealized capital gains (losses), or capital gains overhang (henceforth CGO), for each stock. The CGO of each stock is defined as the difference between the current price and the aggregate reference price, which is defined as the turnover weighted average of past purchase prices using the algorithm defined in Grinblatt and Han (2005).[10] Within each CGO quintile, we construct zero investment portfolios consisting of a long position in high abnormal volume stocks and a short position in low abnormal volume stocks. We take these positions from the first trading day of the next month after the earnings announcement and hold them until the end of the next earnings announcement month or until four months elapse, whichever comes first.

Our novel findings can be summarized as follows.

First, we find that the EAVP is concentrated in CGO quintile 1 (large unrealized losses) and CGO quintile 5 (large unrealized gains). The EAVP in CGO quintile 1 amounts to 9.60% per year, while that in CGO quintile 5 is smaller but still positive and significant at 5.76% per year. The concentration of the EAVP in these quintiles generates

[10] See Section 2.2 for details of how the CGO measure is constructed.

a U-shaped pattern which is robust to various risk adjustments including Fama and French's 3-factor model and its extensions which include price momentum (Carhart (1997)), liquidity (Sadka (2006)) and/or SUE factors.[11] The EAVPs in other quintiles are notably smaller or become insignificant after risk adjustments. This finding suggests that the EAVP is associated with active selling which is not driven by the fundamentals of the underlying stocks but by the magnitude of CGOs. This puts temporary downward pressure on stock prices, which is subsequently corrected. This finding suggests that only those high abnormal trading volumes with large unrealized capital gains or losses will be associated with higher returns.

Second, the EAVP is not merely reflecting a contemporaneous positive relationship between trading volume and high returns (Karpoff (1987)) nor is it short lived. If earnings announcements are uniformly distributed within a month, the average implementation lag of our strategy would be two weeks after an earnings announcement. [12] Thus, on average, the EAVP does not reflect an immediate price increase accompanying abnormal trading volume. However, it is still possible that a significant portion of the EAVP can reflect price changes close to earnings announcements. To examine this issue, we further remove the first month of our original holding period from each portfolio. Thus, for example, if a stock announces earnings in April, it will first enter a portfolio on the first day of June. Even with this modification, we still observe a positive and significant EAVP in the full sample and in both CGO quintiles 1 and 5. Such persistent effects show that our findings are distinct from research that examines the relationship between abnormal volume and returns on the (expected) month of earnings announcement such as Lamont and Frazzini (2007). Existing literature on selling pressure focuses mostly on relatively short term effects. Even so, our longer term effect of selling pressure induced by prior gains or losses is consistent with Frazzini (2006), who finds that the effect of prior gains or losses on the PEAD is quite persistent. More generally, our finding is consistent with Garfinkel

[11] Grinblatt and Han (2005) show that the CGO effect subsumes the momentum effect. After controlling for the CGO effect, momentum loses its predictive power in the cross-sectional regression of stock returns. We find that our CGO effect is mostly independent of the momentum effect and possesses stronger predictive power for the EAVP. See Section 2.4 for details.

[12] Earnings announcements skew slightly towards the end of the month. Mean (median) announcement day is 18.11 (19.00), with a standard deviation of 7.93 days. Mean (median) lag until being included in a portfolio is 14.42 (13.00) days.

and Sokobin (2006), and Lerman *et al.* (2008) who find that abnormal trading volume around earnings announcements has a prolonged impact on PEAD (extending to 60 trading days or one day after the subsequent earnings announcement, respectively).

Third, when portfolios are subdivided by SUE, we find that the EAVP is largest in those cases where stocks with large capital losses receive good news. For these stocks, the EAVP amounts to 10.56% per year. This pattern suggests that good news triggers loss realization, and is consistent with Barber *et al.* (2007) who find that the proportion of losses which are realized increases when the general market appreciates. This may be because investors find it easier to accept losses if an event occurs that lets them recoup even a small part of the losses.[13] This result also shows how the EAVP and PEAD may interact with each other in generating future return patterns. We should note that this behavior of sellers is sub-optimal since they are selling stocks with good news, and such stocks are known to exhibit persistent price increases (Bernard and Thomas (1989, 1990)).

Fourth, the EAVP is observed in small and medium firms (the first and second size terciles) but not in large firms (the third size tercile). However, the sources for the EAVP within small and medium firm subsample are notably different. For small firms, the CGO effect is strongest (11.16% per year) for large unrealized losses (as in the full sample). However, for medium sized firms, the CGO effect is only observed for large unrealized gains. In large firms, the EAVP becomes insignificant even in the two tails of the CGO distribution. This finding is consistent with Gervais *et al.* (2001) and Lerman *et al.* (2008) who find a stronger volume premium for smaller firms. This finding is also consistent with the general empirical pattern that anomalous return behavior is stronger in smaller stocks where individual investors, rather than institutional investors, are major players.

Fifth, the EAVP for stocks with large capital losses is not merely a reflection of the well known January effect. If tax considerations are the main motives for realizing losses, we expect a significant drop in the EAVP when we remove December and January returns since tax-loss selling is concentrated at the end of the year (Grinblatt and

[13] Good news may also facilitate the realization of losses due to the existence of investors who buy on good news to exploit post-earnings announcement drift.

Keloharju (2001, 2004)). However, the EAVPs are virtually the same when we remove these two months.

Finally, all of the results from the calendar time analyses hold for event time analyses and also pass other robustness checks including different sorting mechanism (quintile vs. tercile, absolute and relative cutoffs in sorting stocks) and sub-period analyses.

Our results supplement existing explanations for the EAVP which focus mainly on the motivation of buyers and the resulting upward price pressure. According to Gervais *et al.* (2001), heightened visibility created by an abnormal volume shock broadens a stock's investor base, resulting in upward pressure on the stock price. Barber and Odean (2008), Lamont and Frazzini (2007) and Lerman *et al.* (2008) also stress the buying pressure for attention grabbing stocks in explaining the high volume-high return relationship. These explanations focus on the buy side, given that a volume shock is observed. We supplement this story by examining a possible cause for the high abnormal volume itself. We provide explanations for who become sellers around such attention grabbing events, on what they base their selling decisions, and what the implications of such decisions are for future returns.

Our findings are distinct from recent research that emphasizes the link between information uncertainty and future returns such as Zhang (2006) and Francis *et al.* (2007). Zhang (2006) shows that price momentum is stronger among stocks with more uncertainty. However, Zhang (2006) and Francis *et al.* (2007) discuss their findings based on models of representative investors, such as those of Daniel *et al.* (1998) and Brav and Heaton (2002), which are devoid of trading volume. In addition, increased uncertainty does not necessarily lead to high abnormal volume. Rather, a high degree of uncertainty may prevent investors from making any trades at all until such uncertainty is resolved.

The remainder of the paper is structured as follows. Section 2.2 describes the data. Section 2.3 reports our main results. Section 2.4 discusses various robustness checks and Section 2.5 concludes.

2.2 DATA AND VARIABLE CONSTRUCTION

Our sample consists of 51,175 quarterly earnings announcements made between April 1983 and September 2001, as reported in the Compustat quarterly data file.[14] In keeping with the conventions of prior volume related studies, we include only common shares of NYSE/AMEX companies. This section details the construction methodologies for the main variables used in the chapter.

Abnormal Volume

Our chapter's primary goal is to identify a possible source for the predictive power of volume triggered by earnings announcements. To measure this volume properly, we need to control for the normal level of trading volume for each company (i.e. expected volume were it not an earnings announcement day). Following Tkac (1996) and Lo and Wang (2000), we estimate the normal level of volume by running a market model regression using daily turnover data for the prior calendar year (i.e. *y-1*):

$$TO_{i,t} = \alpha_{i,y-1} + \beta_{i,y-1} \cdot MKTTO_t + e_{i,t} \qquad [1]$$

where $TO_{i,t}$ is the turnover measure for company i on day t (in year *y-1*) and $MKTTO_t$ is the value weighted turnover for the entire market measured on day t (in year *y-1*). The resultant α and β coefficients for company i in year *y-1* are then used to calculate estimated daily turnovers (ESTTO) for company i in year y. Specifically, ESTTO is calculated as:

$$ESTTO_{i,t,y} = \hat{\alpha}_{i,y-1} + \hat{\beta}_{i,y-1} \cdot MKTTO_{t,y} \qquad [2]$$

[14] Quarterly announcement data is consistently available from Compustat starting in the first quarter of 1972. However, since the construction of CGO requires five years of past data, our sample period effectively begins in 1977. The liquidity factor of Sadka (2006) is only available from 1983, and as such further restricts the start of our sample period. However, our results are not affected by the exclusion of the 1977 to 1982 data.

where $ESTTO_{i,t,y}$ is the estimated turnover for stock i on day t of year y and $\hat{\alpha}_{i,y-1}$ and $\hat{\beta}_{i,y-1}$ are the α and β parameter estimates from [1]. The difference between the actual daily turnover and the estimated daily turnover is the market-adjusted volume for the day. Finally, we define abnormal volume for an earnings announcement made on day t as the sum of daily market-adjusted volume over the three day window [*t-1, t+1*].

Table 2.1 shows market-adjusted volume around event time for [*t-5, t+1*]. There is a surge in market-adjusted volume at *t*, consistent with that reported in Lee *et al.* (1993).[15] Positive and significant market-adjusted volume is observed over the three day window [*t-1, t+1*] regardless of whether companies receive good or bad news, or have high or low CGO.[16] Table 2.1 confirms that earnings announcements are attention grabbing events in which active portfolio rebalancing happens among investors.

[Figure 2.1 about here]

Capital Gains Overhang (CGO)

The CGO measure estimates unrealized capital gains or losses. We estimate aggregate CGO around earnings announcements for each company based on the recursive formula as defined in Grinblatt and Han (2005). The basic idea of their methodology is to calculate CGO as the difference between the current stock price and the price at which an average investor would have purchased the stock (i.e. the average reference price). While the current stock price is easily observable, the reference price must be continually recalculated in response to trading activity. For example, a stock price increase with *high* volume means that a large number of shares are purchased at the new (higher) price, and necessitates that the reference price be adjusted up towards the new purchase price (thereby reducing the average capital gain). In contrast, a stock price increase with *low*

15 A detailed description of the characteristics of abnormal volume can be found in Section 2.3.

16 Good (bad) news is defined as the top (bottom) SUE quintile. High (low) CGO is defined as the top (bottom) CGO quintile. Formal definitions of SUE and CGO are given in the following paragraphs.

volume means that fewer shares are purchased at the new (higher) price, and does not necessitate that the average reference price be adjusted as much (thereby increasing the average capital gain).

Grinblatt and Han (2005) use weekly miniCRSP data over a five year horizon to calculate CGO. We also use a five year horizon, but use daily CRSP data in our calculations. Daily frequency allows us to use a consistent interval between earnings announcement dates and the CGO construction window for each stock (i.e. a five year window ending exactly five trading days prior to each earnings announcement, as described below). However, our results are robust to whether we use daily or weekly data in calculating CGO.

For each stock, we first calculate a daily reference price, which represents the volume weighted average purchase price for the stock. To achieve this, we iteratively apply the following reference price formula starting with the first available observation in the CRSP daily data file:

$$R_{t+1} = V_t \cdot P_t + (1 - V_t) \cdot R_t \qquad [3]$$

where t is a trading day, V is the daily turnover,[17] P is the stock price,[18] and R is the reference price. The initial reference price of a company is defined to be the first stock price available in CRSP.[19] Finally, we define the CGO of a company which makes an earnings announcement on day t as follows:

$$CGO_t = \frac{(P_{t-5} - R_{t-5})}{P_{t-5}} \qquad [4]$$

17 Daily turnover is defined as the number of shares traded divided by the total number of shares outstanding.

18 Stock price is calculated as the day's closing price (or the average of bid / ask prices if no closing price is available) divided by the cumulative price adjustment factor in the CRSP daily data file. This is done to correctly account for stock splits and stock dividends.

19 This is the same as assuming that 100% of the shares outstanding have traded on that day. To mitigate the impact of an arbitrarily chosen initial reference price, we use a calibration period of 1,300 trading days (approximately five years) as in Grinblatt and Han (2005). If 1,300 trading days of data is not available for a company prior to the start of our sample period we delay including it in the analyses until its calibration period has completed.

Our main hypothesis is that capital gains or losses *cause* trading at the time of earnings announcements. Thus, if CGO is measured very close to an earnings announcement, it may be contaminated by trading based upon the upcoming earnings announcement itself. In such situations, it is difficult to evaluate the causal relationship between CGO and the event volume it triggers. We try to minimize this concern by lagging CGO values by five trading days.

As in Grinblatt and Han (2005), CGO is positively correlated with market capitalization and past momentum in our sample. It is also negatively correlated with average share turnover in the prior year.[20] Thus, in examining the effect of CGO, controlling for these variables will be important. We will return to this issue in the following section when we implement our trading strategy which utilizes CGO values.

Standardized Unexpected Earnings (SUE)

For those investors who make investment decisions based on capital gains or losses as measured by CGO, a large surprise at an earnings announcement could act as a catalyst to realize these gains or losses. We measure surprise using standardized unexpected earnings (SUE) based on a seasonal random walk hypothesis, where unexpected earnings are calculated as earnings per share for the current quarter less earnings per share for the same quarter, one year prior. We then normalize this difference by dividing it by the standard deviation of the past 20 unexpected earnings values (i.e. five years of data).[21]

Table 2.2 summarizes the day ranges used in the construction of the main variables. In the following section, we implement calendar time trading strategies based on these key variables.

[Figure 2.2 about here]

[20] Correlation results are available from the authors upon request.

[21] If more than 10 of the past 20 unexpected earnings values are missing or invalid, we do not calculate the standard deviation and consider the quarter's SUE value to be missing for the company.

2.3 EMPIRICAL FINDINGS

2.3.1 The Earnings Announcement Volume Premium (EAVP)

We examine the EAVP by constructing calendar time portfolios. We collect all of the earnings announcements in a given quarter and sort them into quintiles based on abnormal volume. A stock is assigned to an abnormal volume quintile portfolio at the start of the next month after the earnings announcement, and is held within that portfolio until the end of the next earnings announcement month or until four months elapse, whichever comes first.[22] All cutoff values are based on the prior quarter's distribution.

If an earnings announcement is made in the first week of a month, the stock will not enter a portfolio until almost four weeks after the earnings announcement. If the earnings announcement dates are uniformly distributed within a month, average implementation lag would be about two weeks.[23] If the high returns of high abnormal volume stocks are concentrated only around earnings announcements (for example, they persist for only one or two weeks immediately after the earnings announcements), our portfolio strategy underestimates the magnitude of abnormal volume's effect on future returns.[24] However, we introduce this lag to measure the persistent impact of high abnormal volume on future returns and to ensure that portfolio rebalancing occurs monthly. With our methodology, each quintile contains an average of approximately 129 to 139 stocks.

[Table 2.1 about here]

Panel A of Table 2.1 shows summary statistics for each abnormal volume quintile. Several interesting patterns emerge.

---

[22] Since abnormal volume is calculated using a [*t-1, t+1*] window around earnings announcements, the abnormal volume of a company which announces earnings on the last trading day of the month requires information from the first trading day of the following month. In such cases, we skip one additional month before including the stock in a portfolio in order to maintain the implementability of our strategy. Such observations represent only 2.7% of the total sample and do not change our results in any significant way.

[23] As discussed in endnote 11, actual mean (median) lag until being included in a portfolio is 14.42 (13.00) days.

[24] In a following subsection we employ an even more conservative implementation lag in order to further examine if the EAVP mainly reflects a contemporaneous positive relationship between returns and abnormal volume.

First, abnormal volume values are quite dispersed. Mean (median) values for abnormal volume quintiles 1 and 5 are -0.0066 (-0.0058) and 0.0188 (0.0182) respectively. Thus, even though earnings announcements are attention grabbing events, there are substantial variations in the resulting abnormal volume.

Second, SUE is monotonically increasing across abnormal volume quintiles, ranging from a mean (median) of -0.1835 (-0.1534) to -0.0026 (0.0508).[25] This suggests that good news tends to generate higher abnormal volume than bad news. This is interesting since analysts' forecast dispersion, which is used as a measure of opinion dispersion (Diether *et al.* (2002)) and is presumed to be a source of trading volume in some research (Ajinkya *et al.* (1991) and Kandel and Pearson (1995)), is reported to be larger in companies with poor earnings (Ciccone (2001)). There is little discussion in the literature as to why good news tends to generate higher abnormal volume than bad news, and exploring possible explanations is one of the topics of this chapter.

Third, the average level of daily turnover in the prior year is higher in abnormal volume quintiles 1 and 5 than in the middle quintiles. This shows that the EAVP cannot be solely explained by different liquidity levels for high and low abnormal volume stocks. Also, Chordia *et al.* (2007a) show that the most illiquid stocks tend not to experience high volume around earnings announcements regardless of whether good or bad news is received. Thus, it seems unlikely that the EAVP is systematically related to the illiquidity premium (Amihud (2002)). Even so, we still control for liquidity related return premiums when examining the returns of our calendar time portfolios.

Fourth, company size displays a hump-shaped pattern, with the smallest values found in abnormal volume quintiles 1 and 5. However, the difference in this variable between abnormal volume quintiles 1 and 5 is significant, so we will include a control for size in our subsequent analyses.

Fifth, B/M values are decreasing across abnormal volume quintiles, and the difference in this variable between abnormal volume quintiles 1 and 5 is significant. Thus, B/M will also be controlled for when examining our calendar time portfolio returns.

[25] Mean SUE values are strongly influenced by a small number of extremely large outliers, especially in the negative tail of the distribution. Winsorizing SUE values at 1% from both tails results in mean values ranging from -0.1437 in abnormal volume quintile 1 to 0.0588 in abnormal volume quintile 5.

Finally, high abnormal volume stocks tend to show higher past returns as measured by 12-month momentum ending immediately prior to the earnings announcement month. This may reflect the fact that SUE and momentum are positively correlated (Chordia and Shivakumar (2006)).

Panel B of Table 2.1 shows summary statistics for each abnormal volume tercile. The same general patterns as in Panel A of Table 2.1 are observed. The reason for reproducing the results of Panel A of Table 2.1 using terciles is as follows. In the following analyses, we will refine our strategy by simultaneously using multiple sorting variables. In such cases, maintaining a sufficient number of stocks in each of the portfolios is important to ensure that our results are not driven by outliers. For this reason, we will use abnormal volume terciles in defining our high minus low portfolios in the refined strategies where multiple sorting variables are involved. Panel A (B) of Table 2.1 shows that each volume quintile (tercile) portfolio in the base-case strategy contains an average of approximately 130 (220) stocks.

Panels C and D of Table 2.1 shows raw and risk-adjusted returns for each abnormal volume quintile and volume tertile portfolio and for a high minus low portfolio. The EAVP is defined as the monthly return to a zero investment portfolio which takes a long position in abnormal volume quintile 5 or abnormal volume tercile 3 (high volume) and a short position in abnormal volume quintile 1 or abnormal volume tercile 1 (low volume), where stocks are equally weighted within each portfolio.

Panel C of Table 2.1 reports the EAVP defined using abnormal volume quintiles. The first column reports raw returns. They are monotonically increasing across abnormal volume quintiles. The EAVP is 0.47% per month (5.64% per year) and is significant.

We also use various benchmark models in order to control for the potentially different risk characteristics of the abnormal volume quintiles. We start by calculating Jensen's alpha using the conventional Fama-French 3-factor model, and find that the magnitude of the risk-adjusted EAVP (0.50% per month (6.00% per year)) is similar to that based on raw returns. Next, we add a momentum factor as discussed in Carhart (1997). The EAVP from this 4-factor model is smaller at 0.32% per month (3.84% per year), but is still significant. Thus, even though the EAVP has some systematic exposure to the momentum factor, it cannot be entirely explained by the momentum effect.

As shown in Panels A and B of Table 2.1, high abnormal volume stocks have higher average SUE values than low abnormal volume stocks. This raises the possibility that the EAVP is related to PEAD, in which high SUE stocks outperform low SUE stocks (Bernard and Thomas (1989, 1990)). To control for PEAD, we incorporate a SUE factor that is defined as the difference in returns between SUE decile 10 (large positive surprise) and SUE decile 1 (large negative surprise) portfolios. The EAVP estimated under the Fama-French 3-factor model augmented by this SUE factor[26] is 0.32% per month (3.84% per year) and remains significant. This shows that the EAVP after earnings announcements is not a mere manifestation of PEAD.

Finally, as a control for any difference in liquidity levels between high and low abnormal volume stocks, we use the liquidity factor as discussed in Sadka (2006), which is designed to capture market-wide liquidity shocks that are not easily diversifiable.[27] Adding this new factor to Carhart's 4-factor model does not change the EAVP in any significant way.

Panel D of Table 2.1 reports results based on abnormal volume terciles. These results are very similar to those based on abnormal volume quintiles.

In summary, the EAVP remains positive and significant even after controlling for the conventional risk factors used in asset pricing literature. Our finding is related to that of Lamont and Frazzini (2007), who find that stocks with high predicted abnormal volume increases in months where earnings announcements are expected tend to have higher returns than stocks with low predicted abnormal volume increases during the same months. However, Lamont and Frazzini (2007) focus on the *contemporaneous* relationship between predicted abnormal volume and returns for the expected month of earnings announcements. A stock goes into the long position from the first day of the month of expected earnings announcement and stays in the portfolio for the month. On

---

[26] We do not include Carhart's momentum factor and the SUE factor at the same time due to their positive correlation (Chordia and Shivakumar (2006)).

[27] Sadka (2006) separates price impact into permanent variable and transitory fixed price components and finds that only the market-wide variation of the permanent variable component is priced. We use this liquidity factor in our tests.

the contrary, our primary focus is on the impact of abnormal volume around earnings announcements on *subsequent* returns. In our case, a stock is assigned to a long position from the first day of the next month after the earnings announcement is made and held until the end of the next earnings announcement month or for 4 months, whichever comes first. Thus, if earnings announcements are evenly distributed across the month, the average lag before a stock is assigned to a portfolio would be 15 days. In our data, mean (median) lag before a stock is included in a portfolio is 14.42 (13.00) days after an earnings announcement. This lag suggests that our results may not be driven by price changes immediately following earnings announcements. In a following subsection we will further examine whether the EAVP is a persistent phenomenon and not simply concentrated in the short period surrounding earnings announcements by implementing an even more conservative lag structure.

In the next subsection, we explore possible sources of the EAVP by examining its relationship with CGO.

2.3.2 The EAVP across CGO quintiles

Our main hypothesis is that the EAVP arises from the interaction between a class of investors who make their selling decisions based on capital gains or losses and arbitrageurs or liquidity providers who exploit the opportunities created by the abnormal selling pressure coming from such investors. To test this hypothesis, we re-examine the trading strategy discussed in the previous subsection in relation to our CGO measure. A stock is assigned to a CGO quintile portfolio at the start of the next month after the earnings announcement based on cutoff values obtained from the prior quarter's CGO distribution. Within each CGO quintile, we construct abnormal volume terciles in a similar manner. Each portfolio contains an average of approximately 45 stocks per month. This additional level of refinement results in 15 (CGO quintile-abnormal volume tercile) portfolios in total.[28]

[28] Our results for this table, and other tables, are similar if we calculate abnormal volume terciles across all observations instead of within each CGO quintile or tercile. See Section 2.4 for more details.

[Table 2.2 about here]

Panel A of Table 2.2 reports the mean and median values for CGO, abnormal volume and number of observations in each portfolio. There is a substantial dispersion in mean (median) CGO values across the CGO quintiles, ranging from -52.00% (-32.14%) for quintile 1 to 37.58% (35.91%) for quintile 5. Mean (median) abnormal volume for the high abnormal volume portfolios ranges from a low of 0.0118 (0.0063) to a high of 0.0137 (0.0076) across the CGO quintiles. These numbers suggest that high abnormal volume is not confined to a particular CGO quintile

Panel B of Table 2.2 reports one of the central findings of the chapter. The EAVP exhibits a strikingly robust U-shaped pattern across CGO quintiles. The raw return EAVP for CGO quintiles 1 and 5 is positive and significant at 0.80% and 0.48% per month (9.60% and 5.76% per year), respectively. The risk-adjusted EAVP remains positive and significant for these two quintiles as well. However, in the middle CGO quintiles (2 through 4), the EAVP has substantially lower magnitude and significance than in the two extreme CGO quintiles, especially with risk adjustments. This pattern suggests that *not all trading volumes are created equal* in their implications for future returns. The abnormal volume around earnings announcements contains information for future returns only in stocks with extreme capital gains or losses.

Another intriguing fact is that the EAVP in CGO quintile 1 is from 1.6 to 2.1 times higher than that in CGO quintile 5. For CGO quintile 1, the risk-adjusted EAVP ranges from a low of 0.74% to a high of 0.81% per month (8.88% to 9.72% per year). For CGO quintile 5 the values range from a low of 0.36% to a high of 0.48% per month (4.32% to 5.76% per year). The EAVP in CGO quintile 5 is consistent with the disposition effect. When investors subject to the disposition effect want to realize gains in past winner stocks, arbitrageurs will see a profit opportunity and absorb the selling pressure. Therefore, we expect to see positive returns following these types of trades, and this is indeed the pattern we find in CGO quintile 5.

It is also important to note that the EAVP is driven mostly by large (abnormal) positive returns to the long (high abnormal volume) positions rather than by large (abnormal) negative returns to the short (low abnormal volume) positions. This shows that the EAVP of CGO quintile 1 is not entirely consistent with the disposition effect, which predicts that investors will be reluctant to realize losses and this lack of selling pressure will drive abnormally low returns for past losers. This implies that the EAVP in CGO quintile 1 should be driven by abnormally low returns to the short (low abnormal volume) positions. However, as in CGO quintile 5, the EAVP in CGO quintile 1 is also being driven by the high returns to the long (high abnormal volume) positions. In fact, the returns of low abnormal volume stocks in CGO quintile 1 are not significantly different from zero in all risk-adjusted models except the 3-factor model.

The large EAVP in CGO quintile 1 suggests that loss realization may play an important role in the relationship between abnormal volume and returns around earnings announcements. Even though loss realization is not extensively discussed in disposition effect related literature, recent research by Grinblatt and Keloharju (2001, 2004), Grinblatt and Moskowitz (2004), and Jin (2006) suggests that it can be an important factor in the cross-section of returns. We further discuss the issue of loss realization in a following subsection.

2.3.3 Persistence of abnormal volume effects

Our portfolio strategy is constructed to incorporate stocks into portfolios at the start of the next month after the earnings announcement. Thus, stocks for companies which make earnings announcements near the end of a month could enter a portfolio as few as two days later, and as such it is still possible that the contemporaneous positive relationship between abnormal volume and returns could explain a significant portion of the EAVP.

To help answer this, we insert an additional lag when forming our portfolios. In other words, a stock that announces its earnings in month m enters a portfolio at the start of month $m+2$, not at the start of month $m+1$. If abnormal volume's impact on returns is concentrated in the short period surrounding announcements, the EAVP should be significantly reduced. We perform this exercise for the base EAVP cases (Panels A and B of Table 2.3) and for the EAVP across CGO quintiles (Panel C of Table 2.3).

[Table 2.3 about here]

Panels A and B of Table 2.3 show that even though the EAVPs unconditioned on CGO are reduced in magnitude (from 0.47% to 0.39% per month (5.64% to 4.68% per year) in volume quintile based EAVP and from 0.40% to 0.33% per month (4.80% to 3.96% per year) in volume tercile based EAVP), they still remain positive and significant in most cases. Panel C of Table 2.3 also shows that the EAVP in CGO quintiles 1 and 5 is largely robust to the lag adjustment. For CGO quintile 1, the raw return EAVP becomes 0.47% per month (5.64% per year) versus 0.80% per month (9.60% per year) without the lag adjustment. Thus, there is still a considerable EAVP even for this conservative portfolio. The risk-adjusted EAVPs in CGO quintile 1 are also significant, with the only exception being the case of the 3-factor model augmented with the SUE factor.

Interestingly, the EAVP in CGO quintile 5 becomes larger and more significant when we skip an additional month. For example, the raw return EAVP becomes 0.62% per month (7.44% per year), which is larger than the 0.48% per month (5.76% per year) observed when we do not omit the first month. Additionally, the EAVP for the 3-factor model augmented with the SUE factor remains highly significant in CGO quintile 5, unlike the unconditioned EAVP shown in Panels A and B of Table 2.3.

Table 2.3 compares our prior findings, and shows that the U-shaped pattern of the EAVP is strongly preserved when we skip an additional month when forming our portfolios.

[Figure 2.3 about here]

In summary, this subsection shows that the EAVP among stocks with large capital gains or losses is quite persistent and is not a mere reflection of the contemporaneous positive relationship between abnormal volume and returns around earnings announcements.

2.3.4 *The impact of news on the EAVP*

So far, we have focused on the relationship between the EAVP and CGO *prior* to the arrival of news. It is possible that investors' final trading decisions depend on both prior CGO and how the current news moves the stock price. For example, Frazzini (2006) argues that post-event drift is larger when the news and CGO have the same sign.

In a prior subsection, we find that the EAVP is mainly concentrated in CGO quintiles 1 and 5. The strong EAVP in quintile 5 is consistent with the existence of selling pressure due to investors' premature realization of potential winners with large capital gains. However, we need a different explanation for what triggers high abnormal volume in stocks with large capital losses and why such volume predicts high subsequent returns. One possible hypothesis is that investors, who behave as the disposition effect posits within the realm of moderate gains or losses, will eventually decide to realize their losses when they exceed a certain threshold. The resulting selling pressure, not related to the stock's fundamentals, will be absorbed by arbitrageurs or liquidity providers, and will lead to higher subsequent returns.

Our hypothesis on loss realization is not entirely new in the literature. For example, when investors hold undiversified portfolios, a large loss in any one stock would imply a substantial decrease in their wealth. Barberis and Huang (2001), Barberis *et al.* (2001), and Gomes (2005) suggest a decrease in wealth may increase investor's risk aversion and thus let the investor opt to realize those losses rather than holding on to them, which puts downward pressure on prices.[29] In addition, for tax sensitive investors, the benefits of realizing losses from a tax standpoint may become too large to ignore and this tax-loss

[29] The loss realization we discuss in this paper is different from popular suggestions among practitioners about stop loss selling. For example, O'Neil (1995) recommends placing a stop loss at -8%. The average capital loss in CGO quintile 1 is -52%, and therefore the loss realization in CGO quintile 1 does not fit well with a typical stop loss strategy. In this regard, investors are reluctant to realize losses until they become extremely large.

selling may have future return implications as in Grinblatt and Keloharju (2001, 2004), Grinblatt and Moskowitz (2004), and Jin (2006).

Next, we ask under what circumstances would investors prone to the disposition effect find it least painful to realize losses. Barber *et al.* (2007) provide an interesting hypothesis on this issue. They find that the proportion of losses realized by Taiwanese investors increases with market returns and interpret this finding as suggesting that investors are more willing to realize losses when they can recoup even a small part of them. If their conjecture is correct, good news may provide investors with an exit opportunity by allowing them to recover a part of their losses. If this is the case, investors would more actively sell stocks with large capital losses after those companies receive good news than bad news. Consequently, we should find that the EAVP in stocks with large prior losses will be concentrated in those which receive good news.

On the other hand, it is also possible that the arrival of additional bad news after investors have already experienced massive losses triggers panic selling, thereby lowering the stock price to a level not warranted by the content of the news. Such a scenario suggests that the EAVP among stocks with large capital losses will be concentrated in those stocks which receive bad news. By empirically examining the impact of news on the EAVP, we may be able to determine which scenario is more plausible.

In order to study the interaction effect of CGO and current news, we examine the EAVP in CGO-SUE double sorted portfolios. A stock is assigned to a SUE tercile portfolio at the start of the next month after the earnings announcement based on cutoff values obtained from the prior quarter's SUE distribution. CGO terciles are defined in the same way as CGO quintiles are defined in previous subsections, and we further divide each CGO tercile into abnormal volume terciles. This results in 27 (SUE tercile-CGO tercile-abnormal volume tercile) portfolios in total. Each portfolio contains an average of approximately 25 observations per month.

[Table 2.4 about here]

Table 2.4 shows the EAVP for CGO terciles 1 and 3 for each SUE tercile. Several interesting patterns emerge.

First, CGO tercile 1-SUE tercile 3 (large capital losses with good news) has the largest raw return EAVP at 0.88% per month (10.56% per year). The risk-adjusted EAVP for this combination is even higher, ranging from 0.97% to 1.07% per month (11.64% to 12.84% per year) and remains highly significant. Examining the risk-adjusted returns of the high and low abnormal volume portfolios separately, we can see that the EAVP for this combination is being driven primarily by large positive returns to the long (high abnormal volume) positions rather than by large negative returns to the short (low abnormal volume) positions. This shows that good news triggers loss realization and that this is the major source of the EAVP.

Second, the raw return EAVP among stocks with large capital losses which receive bad news (CGO tercile 1-SUE tercile 1) is also positive and significant, even though the magnitude is smaller at 0.70% per month (8.40% per year). The risk-adjusted EAVP for this combination ranges from 0.50% to 0.60% per month (6.00% to 7.20% per year) and remains significant. However, unlike the case of stocks with large capital losses with good news, the EAVP for this group is generally not driven by large positive returns to the long (high abnormal volume) positions, especially with risk adjustments.

Third, for CGO tercile 3, the raw return EAVP is positive and significant for SUE terciles 2 and 3 (0.56% and 0.52% per month (6.72% and 6.24% per year), respectively), but negative and insignificant for SUE tercile 1. A similar pattern is found for the risk-adjusted EAVP values. The positive EAVP values are driven mainly by the positive returns to the long (high abnormal volume) positions, especially for SUE tercile 3. We take this as evidence that investors more actively sell winner stocks when the earnings news is not bad, which is consistent with Frazzini (2006).

2.3.5 Firm size

Even though our results survive risk adjustments which include Fama-French's size factor (which is defined as the returns to the first decile (small) stocks minus the returns to the tenth decile (large) stocks), it is still possible that our results may exhibit variations across firm size groups. For example, smaller stocks are more likely to be held by individual investors who could be more susceptible to sub-optimal investment behaviors than institutional investors. If this is the case, we expect stronger CGO effects in smaller stocks. To examine this issue, we first subset our results by size terciles. Size is the market capitalization at the end of the prior calendar year, and size terciles are defined using cutoff values based on the market capitalization distribution of all NYSE companies as of the end of the prior calendar year.[30] Panel A of Table 2.5 reports the EAVP unconditioned on CGO. The EAVP is observed in small and medium firms (size terciles 1 and 2). The EAVP is 0.80% and 0.41% per month (9.60% and 4.92% per year) and highly significant for small and medium firms, respectively. The EAVP for both small and medium firms survive risk adjustments. However, the EAVP is not observed in large firms. Panel B of Table 2.5 reports the EAVP across CGO terciles subsetted by size tercile. In small firms, the EAVP is most strongly observed in CGO tercile 1, with a value of 0.93% per month (11.16% per year), suggesting sub-optimal realization of losses (regardless of a firm's future prospects) plays a significant role in the EAVP for this group. The EAVP for CGO tercile 3 is smaller at 0.72% per month (8.64% per year) but also highly significant. Thus, for small stocks, realizing both losses and gains is the source of the EAVP. On the contrary, the EAVP for medium firms is significant only in CGO tercile 3 at 0.61% per month (7.32% per year). This implies the sub-optimal realization of gains, consistent with the prediction of the disposition effect, is the major source of the EAVP in this group. In large stocks, the EAVP is not observed in any CGO tercile. These results are consistent with the findings of Gervais *et al.* (2001) and Lerman *et al.* (2008) who report that a stronger volume and future return relationship is found among smaller stocks.

[Table 2.5 about here]

[30] Observations with size values below the 35th percentile are assigned to size tercile 1, those between the 35th and 65th percentile are assigned to size tercile 2, and those above the 65th percentile are assigned to size tercile 3.

2.4 ROBUSTNESS CHECKS

In this section, we present additional robustness checks for our results.

2.4.1 Absolute cutoff points

So far in our analyses, we defined volume terciles within each CGO quintile or tercile. Such portfolio formation schemes were used in order to maintain a roughly equal number of stocks in each portfolio. However, under this relative scheme, the distribution of abnormal volume in the high (low) volume portfolio of a particular CGO quintile or tercile may be significantly different from that of the high (low) volume portfolios of other CGO quintiles. Even though the descriptive statistics of abnormal volume shown in Panel A of Table 2.2 mitigate such concerns, we replicated all our analyses using volume terciles defined using absolute cutoff values based on the prior quarter's distribution of abnormal volume unconditioned on CGO values. We confirm that our results do not change in any significant way with this change in portfolio formation scheme.[31]

2.4.2 End of the year and January effects

We check whether our results are driven by the well known end of the year or January effects. Loser stocks (i.e. those with poor past performance and presumed large capital losses) are known to experience heavy selling pressure in December and consequently rebound in January. This pattern is often explained by tax-loss selling. If tax considerations are the main motive for realizing losses, and tax-loss selling is concentrated at the end of the year (Grinblatt and Keloharju (2001, 2004), Grinblatt and Moskowitz (2004), and Starks *et al.* (2006)), our results will show strong seasonality as well. We report the EAVP calculated after excluding December and January in Table 2.6.[32] The U-shaped pattern and the magnitude of the EAVP are virtually the same as

[31] Results are available from the authors upon request.
[32] Excluding either December or January separately does not change our results in any significant way.

those reported for the full sample. This does not imply that taxes are not an important factor in investors' decisions about realizing losses. Rather, it implies that the realization of losses may not be confined to these two months.

[Table 2.6 about here]

2.4.3 Sub-period analyses

We also check whether our results are concentrated in certain time periods. We divide our sample into two sub-periods, the 1980s and 1990s, and run our portfolio analyses separately.[33] The results are shown in Table 2.7. U-shaped patterns are observed in both periods, although the EAVP for CGO quintile 1 is much higher in the 1990s as compared to the 1980s. This could be due to the downward trend in transaction costs (Chordia *et al.* (2007a)), which in turn leads to increased trading. Results for CGO quintile 5 are varied, as the EAVP for raw returns and the 3-factor model both increase slightly in magnitude in the 1990s (and remain significant) while the EAVP for the other three risk-adjusted models decrease in magnitude and significance. Overall, we conclude that our findings are not exclusive to one decade.

[Table 2.7 about here]

2.4.4 Cross-sectional analyses

In this subsection, we check whether the results obtained from our calendar time portfolio analyses hold in event time. In order to examine CGO's incremental effect on the influence of abnormal volume on future returns, we run pooled sample least squares cross-sectional regressions using company-quarter observations. To correct for possible correlation within companies or across companies for a given quarter, we calculate

[33] The inclusion of the last year in our sample (2001) to the 1990s sub-period does not change our results in any significant way.

standard errors by clustering observations either by company or by year-quarter as suggested by Peterson (2007).[34] Our regression models are of the form:

$$y_{i,t} = \beta \cdot x_{i,t} + \alpha_t + \varepsilon_{i,t} \qquad [5]$$

where i is the company, t is the calendar quarter in which the company makes the earnings announcement, and α_t is a year-quarter dummy (i.e. a time fixed effect) which controls for any economy-wide shock. In our data, the latter part of the sample contains more company-quarter observations than the earlier part of the sample. Thus, ordinary least squares (OLS) regressions could be unduly influenced by these latter observations. To address this issue, we run weighted least squares (WLS) regressions by deflating each observation using the total number of observations for a given quarter as in Vuolteenaho (2002).

Our dependent variable, DRIFT60, is calculated as the cumulative sum of daily size- and B/M-adjusted returns from days $t+2$ to $t+61$.[35] Abnormal volume data is sorted into deciles by year-quarter, and this decile number (henceforth VOL) replaces the raw abnormal volume value in the following regressions. This smoothing is motivated by the fact that raw abnormal volume values are residuals from company level regressions, and as such could be noisy.[36] Table 2.8 shows the results of the WLS regressions. Standard errors are adjusted by clustering observations by year-quarter.[37]

[Table 2.8 about here]

Regression 1 is the base case. Here, we regress DRIFT60 on SUE, SUE interacted with the natural logarithm of company size (henceforth LSIZE_SUE), and VOL. SUE

---

34 Since it is impossible to correct for both cross-company and within-company correlations simultaneously, we implement clustering by company and clustering by year-quarter separately.

35 The size- and B/M-matched portfolios are constructed using the methodology of Brav and Gompers (1997), which in turn uses the book value specification of Fama and French (1993).

36 As a robustness check we also run the regressions using raw volume. Doing so does not change our results in any significant way.

37 Clustering observations by company does not change our results in any significant way.

will control for the positive correlation between earnings surprise and subsequent returns, and LSIZE_SUE will control for the documented fact that the magnitude of PEAD is lower in large stocks. Consistent with the previous portfolio analyses, we find that the effect of abnormal volume on DRIFT60 is positive and significant.

Next, we examine the variation in the predictive power of VOL in relation to CGO. For regression 2, we add VOL interacted with CGO (henceforth VOL_CGO) to the base case. The coefficient for this interaction term is significant and negative, implying that the positive impact of abnormal volume on returns is stronger for lower CGO values. This is consistent with the results from our calendar time portfolio analyses which show that the CGO effect is stronger in CGO quintile 1. To further investigate this finding, in regression 3 we replace VOL_CGO with the absolute values of positive and negative CGO separately interacted with VOL (henceforth VOL_CGOPOS and VOL_CGONEG, respectively). Consistent with the U-shaped patterns we identified in the calendar time portfolio analyses, the coefficients of VOL_CGOPOS and VOL_CGONEG are both positive and significant, with VOL_CGONEG having the larger coefficient.

One concern is that the magnitude of CGO may be correlated with the degree of illiquidity. This is because high turnover accelerates the resetting of the reference price towards the current market price, thereby reducing the magnitude of CGO. Thus, the positive impact of absolute CGO on DRIFT60 may just reflect an illiquidity premium. Also, CGO may be correlated with momentum, and thus the impact any CGO related variables have on DRIFT60 may simply be capturing the well known momentum effect. To address these concerns, in regression 4 we include both the natural logarithm of average daily turnover measured over the 52-week period prior to the earnings announcement week and the 12-month price momentum. Additionally, we also include momentum squared interacted with VOL to see if the absolute values of CGO interacted with VOL are capturing any non-linear momentum effects. Adding these new control variables does not change our results in any significant way. Thus, we conclude that the positive impact the absolute value of CGO has on returns is not merely the result of an illiquidity premium or a momentum effect.

Overall, we conclude that our calendar time results remain valid in event time.

2.5 CONCLUSION

In this chapter, we examine a calendar time trading strategy which takes long (short) positions in stocks that experience high (low) abnormal volume around earnings announcements. This strategy generates both statistically and economically significant profits, which we define as the *earnings announcement volume premium* (EAVP). We find that the EAVP is mainly concentrated in stocks with either large capital gains or losses as measured by the aggregate capital gains overhang (CGO) metric developed by Grinblatt and Han (2005). This U-shaped pattern for the EAVP with respect to CGO values is robust to various risk adjustments such as size and B/M (Fama and French (1993)), momentum (Carhart (1997)) and liquidity (Sadka (2006)). The EAVP is also shown to be different from the well known post-earnings announcement drift (Bernard and Thomas (1989, 1990)).

Our results show that the major source of the EAVP is either the premature realization of gains as posited by disposition effect, or the realization of extreme losses, both of which exert excessive downward pressure on stock prices resulting in subsequent price corrections. Our results extend existing analyses on the effect of CGO on price momentum (Grinblatt and Han (2005)) or on PEAD (Frazzini (2006)), and add that the well known disposition effect may not hold for stocks with extreme losses. The findings are consistent with recent theoretical and empirical research which stress loss realization for extreme losers (Barberis and Huang (2001), Barberis *et al.* (2001), Grinblatt and Keloharju (2001, 2004), Grinblatt and Moskowitz (2004), Gomes (2005), Starks *et al.* (2006), and Jin (2006)).

We also show that the EAVP of CGO quintile 1 (unrealized losses) and CGO quintile 5 (unrealized gains) extends well beyond the first month after the earnings announcements. This sets our results apart from Lamont and Frazzini (2007) who analyze the contemporaneous impact of concentrated trading around earnings announcements on returns for the months of earnings announcements. Our findings are not driven by the end of the year effect, and are also observed in decade sub-samples. Finally, the results are also present in event time analyses.

We find that the EAVP is largest among stocks with large capital losses and good earnings news. This implies that investors realize big losses when the arrival of good news provides an opportunity for them to recoup, even partially, past losses. These results are consistent with Barber *et al.* (2007) who find that the percentage of realized losses increases when the overall market goes up and hypothesize that investors find it easier to realize losses when they can recoup even a small portion of those losses. However, the timing of such selling decision turns out to be bad since those stocks receiving good news start performing well.

We also find that the EAVP is strongest in small stocks (first market capitalization tercile). In small stocks, the EAVP is observed in stocks with either large unrealized capital gains or losses, as in the full sample. However, in this group a much stronger EAVP is observed in stocks with large unrealized capital losses, suggesting selling of stocks with extremely large capital losses plays an important role in the EAVP. However, in medium sized stocks, the EAVP is only significant in stocks with large unrealized capital gains. Further, the EAVP is not observed in any CGO subgroups for large stocks. This implies that the concentration of the EAVP on stocks with large unrealized capital gains and losses in the full sample is mainly driven by small and medium firms. This finding is consistent with Gervais *et al.* (2001) and Lerman *et al.* (2008) who find a stronger volume and future return relation among smaller companies. Our findings are consistent with a conjecture that investors who are subject to irrational behavior play a more dominant role in small stocks than in large stocks.

The EAVP is not driven by the well known January effect since the magnitude of the EAVP changes little when we exclude December and January returns.

Our findings supplement existing explanations on high volume and high abnormal return relationships discussed in the literature. Gervais *et al.* (2001), Lamont and Frazzini (2007), and Lerman *et al.* (2008) suggest increased visibility or attention for stocks which experience large volume shocks increases the potential investors base for the stocks, generating upward pressure in the stock prices. This explanation focuses on the buy side given that a volume shock is observed. We supplement this story by examining a possible cause for the high abnormal volume itself, providing explanations on who become the sellers around such attention grabbing events, on what they base their selling decisions, and what the implications of such decisions are for future returns.

The robust empirical findings in this chapter could be further researched at a micro level by using data on individual investors' transactions. This would undoubtedly help us gain a greater understanding of the relationship between abnormal volume and future returns, and help explain why not all trading volumes are created equal.

Table 2.1 - Summary statistics, raw and risk-adjusted returns for monthly calendar time portfolios sorted by abnormal volume

Panel A reports summary statistics for each abnormal volume quintile. For each earnings announcement, we define the earnings announcement window as the three trading day interval, $[t-1, t+1]$, where t is the earnings announcement day. Abnormal volume is defined as the sum of daily market-adjusted volume for the earnings announcement window. Each stock is assigned to an abnormal volume quintile portfolio starting from the next month after the end of the earnings announcement window and ending at the end of the next earnings announcement month, or until four months elapse, whichever comes first. All cutoff values are based on the prior quarter's distribution. SUE is standardized unexpected earnings based on a seasonal random walk model. Turnover is the average daily share turnover over the 52-week period ending immediately prior to the week in which the earnings announcement is made. The capital gains overhang (CGO) of a stock is defined as the difference between the current stock price and the reference price, which is then normalized by the current stock price as defined in Grinblatt and Han (2005). The reference price is recursively calculated using the prior five years of transaction price and turnover data. We measure the CGO of each company as of five trading days before an earnings announcement. Size is the market capitalization (in millions of dollars) as of the end of the prior calendar year. B/M is the book to market ratio as defined in Fama and French (1993). Momentum is compounded returns (inclusive of dividends) for the 12 calendar months ending immediately prior to the month in which the earnings announcement is made. The last row shows differences in these variables between abnormal volume quintiles 5 and 1, along with the associated *p-values*. Results with *p-values* below 0.05 (0.10) are marked with ** (*) and are in bold. Panel B reports similar summary statistics by abnormal volume tercile. Panel C reports raw and risk-adjusted returns (Jensen's alphas) from various factor model specifications for each abnormal volume quintile portfolio and zero investment portfolios (ZIPs). The earnings announcement volume premium (EAVP) is defined as the monthly return to a ZIP which takes a long position in abnormal volume quintile 5 (high volume) and a short position in abnormal volume quintile 1 (low volume). Stocks are equally weighted within each portfolio. In adjusting for risk factors through regressions, dependent variables for abnormal volume quintile portfolios are raw returns minus the risk-free (t-bill) rate, and dependent variables for ZIPs are raw returns for the high abnormal volume portfolios minus raw returns for the low abnormal volume portfolios. 3F regressions use the standard Fama–French three factors; 4F regressions add the momentum factor as defined in Carhart (1997) to the 3F specification; 3F+SUE Factor regressions add a SUE factor to the 3F specification, where the SUE factor is calculated as the difference in monthly equally weighted returns between the highest (decile 10) and the lowest (decile 1) SUE portfolios; 4F+Liq Factor regressions add the permanent variable liquidity factor as defined in Sadka (2006) to the 4F regressions. Results with *p-values* below 0.05 (0.10) are marked with ** (*) and are in bold. Panel D reports similar returns for each abnormal volume tercile portfolio and zero investment portfolios (ZIPs) which take a long position in abnormal volume tercile 3 (high volume) and a short position in abnormal volume tercile 1 (low volume).

Abnormal Volume Quintile / Tercile		Obs per Month	Abnormal Volume	SUE	Turnover	CGO	Size	B/M	Momentum
Panel A: Summary statistics by abnormal volume quintile									
1	Mean	139.4	-0.0066	-0.1835	0.0171	-0.0522	1,974	0.8368	0.1097
	Median	140.0	-0.0058	-0.1534	0.0157	-0.0417	1,660	0.8040	0.0954
2	Mean	130.8	-0.0017	-0.1194	0.0096	0.0134	3,199	0.8093	0.1320
	Median	131.0	-0.0015	-0.1064	0.0091	0.0307	2,786	0.7764	0.1291
3	Mean	128.9	0.0001	-0.0966	0.0092	0.0574	4,124	0.7877	0.1527
	Median	128.0	0.0001	-0.0740	0.0087	0.0763	3,402	0.7687	0.1496
4	Mean	134.8	0.0031	-0.0418	0.0117	0.0525	3,728	0.7719	0.1952
	Median	136.0	0.0030	-0.0502	0.0111	0.0680	2,944	0.7641	0.1917
5	Mean	138.3	0.0188	-0.0026	0.0184	0.0280	2,493	0.7560	0.2926
	Median	139.0	0.0182	0.0508	0.0172	0.0458	2,030	0.7331	0.2962
High–Low			**0.0253****	**0.1809****	**0.0013****	**0.0802****	**519****	**-0.0808****	**0.1830****
p-value			*(0.0000)*	*(0.0000)*	*(0.0027)*	*(0.0000)*	*(0.0004)*	*(0.0000)*	*(0.0000)*

Abnormal Volume Quintile / Tercile		Obs per Month	Abnormal Volume	SUE	Turnover	CGO	Size	B/M	Momentum
Panel B: Summary statistics by abnormal volume tercile									
1	Mean	225.7	-0.0048	-0.1575	0.0143	-0.0315	2,375	0.8293	0.1165
	Median	227.0	-0.0042	-0.1328	0.0135	-0.0125	2,067	0.8068	0.1115
2	Mean	218.7	0.0002	-0.0964	0.0094	0.0518	4,054	0.7848	0.1555
	Median	219.0	0.0003	-0.0942	0.0089	0.0744	3,372	0.7610	0.1508
3	Mean	227.7	0.0129	-0.0133	0.0160	0.0384	2,842	0.7619	0.2591
	Median	228.0	0.0126	-0.0008	0.0151	0.0549	2,250	0.7564	0.2633
High–Low			**0.0177****	**0.1442****	**0.0017****	**0.0699****	**467****	**-0.0674****	**0.1426****
p-value			*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.0030)*	*(0.0000)*	*(0.0000)*

Abnormal Volume Quintile / Tercile	Raw Return	Alpha Values 3F	4F	3F + SUE Factor	4F + Liq Factor
Panel C: Raw and risk adjusted returns by abnormal volume quintile					
1	0.0096	**-0.0032****	-0.0001	**-0.0026***	0.0001
		(0.0263)	*(0.9385)*	*(0.0864)*	*(0.9554)*
2	0.0111	-0.0010	0.0005	-0.0013	0.0009
		(0.3811)	*(0.6433)*	*(0.2890)*	*(0.4102)*
3	0.0125	0.0002	0.0016	0.0001	**0.0020***
		(0.8419)	*(0.1700)*	*(0.9462)*	*(0.0762)*
4	0.0136	0.0011	**0.0026****	0.0009	**0.0031****
		(0.3724)	*(0.0253)*	*(0.4608)*	*(0.0080)*
5	0.0143	0.0017	**0.0032****	0.0006	**0.0033****
		(0.2003)	*(0.0189)*	*(0.6568)*	*(0.0156)*
High–Low	**0.0047****	**0.0050****	**0.0032****	**0.0032****	**0.0032****
p-value	*(0.0005)*	*(0.0002)*	*(0.0098)*	*(0.0141)*	*(0.0114)*
Panel D: Raw and risk adjusted returns by abnormal volume tercile					
1	0.0102	**-0.0024***	0.0002	-0.0020	0.0005
		(0.0594)	*(0.8490)*	*(0.1310)*	*(0.6592)*
2	0.0123	0.0002	0.0016	0.0001	**0.0020***
		(0.8664)	*(0.1322)*	*(0.9571)*	*(0.0528)*
3	0.0142	0.0017	**0.0031****	0.0009	**0.0033****
		(0.1618)	*(0.0094)*	*(0.4838)*	*(0.0055)*
High–Low	**0.0040****	**0.0041****	**0.0029****	**0.0029****	**0.0028****
p-value	*(0.0001)*	*(0.0000)*	*(0.0030)*	*(0.0042)*	*(0.0037)*

Table 2.2 - Summary statistics, raw and risk-adjusted returns for monthly calendar time portfolios double sorted by CGO and abnormal volume

Panel A reports the mean (median) values of CGO percent and abnormal volume by CGO quintile across all abnormal volume terciles (full sample), as well as for the high and low abnormal volume terciles separately and for the difference between the high and low abnormal volume terciles. The average number of monthly observations is also reported across all abnormal volume terciles, as well as for the low and high abnormal volume terciles separately. For each earnings announcement, we define the earnings announcement window as the three trading day interval, [*t-1*, *t+1*], where *t* is the earnings announcement day. The capital gains overhang (CGO) of a stock is defined as the difference between the current stock price and the reference price, which is then normalized by the current stock price as defined in Grinblatt and Han (2005). The reference price is recursively calculated using the prior five years of transaction price and turnover data. We measure the CGO of each company as of five trading days before an earnings announcement. Each stock is assigned to a CGO quintile portfolio starting from the next month after the end of the earnings announcement window and ending at the end of the next earnings announcement month, or until four months elapse, whichever comes first. Abnormal volume is defined as the sum of daily market-adjusted volume for the earnings announcement window. Within each CGO quintile, abnormal volume terciles are assigned in the same manner as CGO quintiles. All cutoff values are based on the prior quarter's distribution. Panel B reports raw and risk-adjusted returns (Jensen's alphas) from various factor model specifications for high and low abnormal volume tercile portfolios and zero investment portfolios (ZIPs) by CGO quintile. The earnings announcement volume premium (EAVP) is defined as the monthly return to a ZIP which takes a long position in abnormal volume tercile 3 (high volume) and an equivalent short position in abnormal volume tercile 1 (low volume). Stocks are equally weighted within each portfolio. In adjusting for risk factors through regressions, dependent variables for high and low abnormal volume tercile portfolios are raw returns minus the risk-free (t-bill) rate, and dependent variables for ZIPs are raw returns for the high abnormal volume portfolios minus raw returns for the low abnormal volume portfolios. 3F regressions use the standard Fama–French three factors; 4F regressions add the momentum factor as defined in Carhart (1997) to the 3F specification; 3F+SUE Factor regressions add a SUE factor to the 3F specification, where the SUE factor is calculated as the difference in monthly equally weighted returns between the highest (decile 10) and the lowest (decile 1) SUE portfolios; 4F+Liq Factor regressions add the permanent variable liquidity factor as defined in Sadka (2006) to the 4F regressions. Results with *p-values* below 0.05 (0.10) are marked with ** (*) and are in bold.

Panel A: Summary statistics by abnormal volume tercile within CGO quintile

Abnormal Volume Tercile		CGO Quintile 1	2	3	4	5
		CGO Percent values				
Full Sample	Mean	**-0.5200****	**-0.0266****	**0.1025****	**0.2080****	**0.3758****
	p-value	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.0000)*
	Median	-0.3214	-0.0103	0.1073	0.2078	0.3591
1	Mean	**-0.4768****	**-0.0254****	**0.1011****	**0.2074****	**0.3654****
	p-value	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.0000)*
	Median	-0.3078	-0.0099	0.1062	0.2062	0.3510
3	Mean	**-0.4924****	**-0.0294****	**0.1022****	**0.2076****	**0.3652****
	p-value	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.0000)*
	Median	-0.3093	-0.0143	0.1056	0.2070	0.3499
High–Low	Mean	**-0.0155***	**-0.0040****	0.0011	0.0003	-0.0002
	p-value	*(0.0849)*	*(0.0154)*	*(0.3629)*	*(0.8120)*	*(0.8809)*
	Median	-0.0016	-0.0044	-0.0007	0.0008	-0.0011
		Abnormal volume values				
Full Sample	Mean	**0.0018****	**0.0027****	**0.0030****	**0.0032****	**0.0033****
	p-value	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.0000)*
	Median	-0.0005	-0.0002	0.0001	0.0003	0.0003
1	Mean	**-0.0064****	**-0.0057****	**-0.0048****	**-0.0040****	**-0.0029****
	p-value	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.0000)*
	Median	-0.0047	-0.0042	-0.0035	-0.0028	-0.0019
3	Mean	**0.0125****	**0.0137****	**0.0129****	**0.0129****	**0.0118****
	p-value	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.0000)*
	Median	0.0075	0.0076	0.0074	0.0076	0.0063
High–Low	Mean	**0.0189****	**0.0193****	**0.0177****	**0.0169****	**0.0147****
	p-value	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.0000)*
	Median	0.0122	0.0118	0.0110	0.0104	0.0082
		Number of observations				
Full Sample	Mean	141.4	131.4	130.4	134.9	134.1
	Median	129.0	130.0	131.0	134.0	132.0
1	Mean	48.7	44.3	42.4	43.6	44.9
	Median	46.0	43.0	41.0	44.0	44.0
3	Mean	47.0	44.4	45.1	46.1	46.8
	Median	43.0	43.0	44.0	44.0	45.0

Panel B: Raw and risk adjusted returns by abnormal volume tercile within CGO quintile

Abnormal Volume Tercile		CGO Quintile 1	2	3	4	5
		Raw returns				
1	Mean	**0.0093****	**0.0097****	**0.0078****	**0.0120****	**0.0116****
	p-value	*(0.0307)*	*(0.0021)*	*(0.0086)*	*(0.0001)*	*(0.0001)*
3	Mean	**0.0173****	**0.0129****	**0.0109****	**0.0126****	**0.0165****
	p-value	*(0.0001)*	*(0.0006)*	*(0.0013)*	*(0.0003)*	*(0.0000)*
High–Low	Mean	**0.0080****	**0.0031***	**0.0031***	0.0006	**0.0048****
	p-value	*(0.0001)*	*(0.0918)*	*(0.0607)*	*(0.7223)*	*(0.0050)*
		3 Factor alpha values				
1	Alpha	**-0.0043***	**-0.0030***	**-0.0042****	-0.0005	0.0005
	p-value	*(0.0833)*	*(0.0733)*	*(0.0083)*	*(0.7143)*	*(0.7455)*
3	Alpha	0.0037	-0.0002	-0.0013	0.0001	**0.0053****
	p-value	*(0.1539)*	*(0.8988)*	*(0.4104)*	*(0.9397)*	*(0.0028)*
High–Low	Alpha	**0.0081****	0.0027	**0.0030***	0.0007	**0.0048****
	p-value	*(0.0002)*	*(0.1238)*	*(0.0611)*	*(0.6917)*	*(0.0038)*
		4 Factor alpha values				
1	Alpha	0.0017	0.0000	**-0.0026***	-0.0002	-0.0006
	p-value	*(0.3743)*	*(0.9911)*	*(0.0989)*	*(0.8815)*	*(0.7094)*
3	Alpha	**0.0091****	0.0018	-0.0007	0.0000	**0.0039****
	p-value	*(0.0001)*	*(0.2956)*	*(0.6436)*	*(0.9946)*	*(0.0305)*
High–Low	Alpha	**0.0074****	0.0018	0.0019	0.0002	**0.0045****
	p-value	*(0.0010)*	*(0.3253)*	*(0.2401)*	*(0.8908)*	*(0.0089)*
		3F + SUE Factor alpha values				
1	Alpha	-0.0012	-0.0020	**-0.0047****	**-0.0025***	-0.0008
	p-value	*(0.6250)*	*(0.2564)*	*(0.0054)*	*(0.0865)*	*(0.6277)*
3	Alpha	**0.0063****	-0.0005	-0.0025	-0.0025	0.0028
	p-value	*(0.0182)*	*(0.7916)*	*(0.1273)*	*(0.1418)*	*(0.1150)*
High–Low	Alpha	**0.0076****	0.0015	0.0022	0.0000	**0.0036****
	p-value	*(0.0009)*	*(0.4274)*	*(0.1757)*	*(0.9794)*	*(0.0359)*
		4F + Liq Factor alpha values				
1	Alpha	0.0018	0.0003	-0.0024	0.0004	-0.0001
	p-value	*(0.3656)*	*(0.8428)*	*(0.1297)*	*(0.7951)*	*(0.9718)*
3	Alpha	**0.0094****	0.0019	-0.0003	0.0005	**0.0042****
	p-value	*(0.0000)*	*(0.2739)*	*(0.8294)*	*(0.7596)*	*(0.0202)*
High–Low	Alpha	**0.0076****	0.0016	0.0021	0.0001	**0.0043****
	p-value	*(0.0009)*	*(0.3800)*	*(0.1978)*	*(0.9322)*	*(0.0139)*

Table 2.3 - Raw and risk-adjusted returns for monthly calendar time portfolios sorted by abnormal volume and monthly calendar time portfolios double sorted by CGO and abnormal volume (month 1 removed)

Panel A reports raw and risk-adjusted returns (Jensen's alphas) from various factor model specifications for each abnormal volume quintile portfolio and zero investment portfolios (ZIPs) after removing the first month of the holding period return of each stock in each portfolio. For each earnings announcement, we define the earnings announcement window as the three trading day interval, [$t-1$, $t+1$], where t is the earnings announcement day. Abnormal volume is defined as the sum of daily market-adjusted volume for the earnings announcement window. Each stock is assigned to an abnormal volume quintile portfolio starting from the month $m+2$, where m is the month of the end of the earnings announcement window, and ending at the end of the next earnings announcement month, or until three months elapse, whichever comes first. All cutoff values are based on the prior quarter's distribution. The earnings announcement volume premium (EAVP) is defined as the monthly return to a ZIP which takes a long position in abnormal volume quintile 5 (high volume) and a short position in abnormal volume quintile 1 (low volume). Stocks are equally weighted within each portfolio. In adjusting for risk factors through regressions, dependent variables for abnormal volume quintile portfolios are raw returns minus the risk-free (t-bill) rate, and dependent variables for ZIPs are raw returns for the high abnormal volume portfolios minus raw returns for the low abnormal volume portfolios. 3F regressions use the standard Fama–French three factors; 4F regressions add the momentum factor as defined in Carhart (1997) to the 3F specification; 3F+SUE Factor regressions add a SUE factor to the 3F specification, where the SUE factor is calculated as the difference in monthly equally weighted returns between the highest (decile 10) and the lowest (decile 1) SUE portfolios; 4F+Liq Factor regressions add the permanent variable liquidity factor as defined in Sadka (2006) to the 4F regressions. Panel B reports similar returns for each abnormal volume tercile portfolio and zero investment portfolios (ZIPs) which take a long position in abnormal volume tercile 3 (high volume) and a short position in abnormal volume tercile 1 (low volume). Panel C reports raw and risk-adjusted returns (Jensen's alphas) for high and low abnormal volume tercile portfolios and zero investment portfolios (ZIPs) by CGO quintile, after removing the first month of the holding period return of each stock in each portfolio. The capital gains overhang (CGO) of a stock is defined as the difference between the current stock price and the reference price, which is then normalized by the current stock price as defined in Grinblatt and Han (2005). The reference price is recursively calculated using the prior five years of transaction price and turnover data. We measure the CGO of each company as of five trading days before an earnings announcement. Each stock is assigned to a CGO quintile portfolio starting from the month $m+2$, where m is the month of the end of the earnings announcement window, and ending at the end of the next earnings announcement month, or until three months elapse, whichever comes first. Within each CGO quintile, abnormal volume terciles are assigned in the same manner as CGO quintiles. All cutoff values are based on the prior quarter's distribution. Results with *p-values* below 0.05 (0.10) are marked with ** (*) and are in bold.

		Alpha Values			
Abnormal Volume Quintile / Tercile	Raw Return	3F	4F	3F + SUE Factor	4F + Liq Factor
Panel A: Raw and risk adjusted returns by abnormal volume quintile					
1	0.0106	-0.0023	0.0005	-0.0018	0.0007
		(0.1739)	*(0.7472)*	*(0.3184)*	*(0.6606)*
2	0.0118	-0.0001	0.0014	-0.0006	0.0017
		(0.9177)	*(0.2750)*	*(0.6706)*	*(0.1676)*
3	0.0122	0.0002	0.0014	0.0000	0.0019
		(0.8957)	*(0.2719)*	*(0.9798)*	*(0.1460)*
4	0.0133	0.0009	**0.0027****	0.0008	**0.0032****
		(0.5112)	*(0.0433)*	*(0.5969)*	*(0.0161)*
5	0.0145	0.0020	**0.0035****	0.0003	**0.0037****
		(0.2150)	*(0.0365)*	*(0.8540)*	*(0.0258)*
High–Low	**0.0039****	**0.0043****	**0.0030***	0.0021	**0.0030***
p-value	*(0.0175)*	*(0.0087)*	*(0.0735)*	*(0.2105)*	*(0.0698)*
Panel B: Raw and risk adjusted returns by abnormal volume tercile					
1	0.0110	-0.0016	0.0008	-0.0013	0.0011
		(0.2684)	*(0.5534)*	*(0.4028)*	*(0.4331)*
2	0.0121	0.0003	0.0016	0.0001	**0.0021***
		(0.8304)	*(0.1525)*	*(0.9615)*	*(0.0666)*
3	0.0143	0.0019	**0.0034****	0.0007	**0.0037****
		(0.1819)	*(0.0166)*	*(0.6478)*	*(0.0088)*
High–Low	**0.0033****	**0.0035****	**0.0026****	0.0020	**0.0026****
p-value	*(0.0066)*	*(0.0035)*	*(0.0335)*	*(0.1061)*	*(0.0308)*

Panel C: Raw and risk adjusted returns by CGO quintile by abnormal volume tercile

Abnormal Volume Tercile		CGO Quintile 1	2	3	4	5
		Raw returns				
1	Mean	**0.0099****	**0.0092****	**0.0085****	**0.0140****	**0.0120****
	p-value	*(0.0351)*	*(0.0053)*	*(0.0117)*	*(0.0000)*	*(0.0003)*
3	Mean	**0.0146****	**0.0135****	**0.0116****	**0.0148****	**0.0182****
	p-value	*(0.0011)*	*(0.0008)*	*(0.0018)*	*(0.0001)*	*(0.0000)*
High–Low	Mean	**0.0047***	**0.0043***	0.0031	0.0008	**0.0062****
	p-value	*(0.0742)*	*(0.0956)*	*(0.1700)*	*(0.7118)*	*(0.0030)*
		3 Factor alpha values				
1	Mean	-0.0043	-0.0033	**-0.0039***	0.0014	0.0007
	p-value	*(0.1573)*	*(0.1063)*	*(0.0591)*	*(0.4372)*	*(0.6827)*
3	Mean	0.0013	0.0003	-0.0004	0.0023	**0.0069****
	p-value	*(0.6671)*	*(0.8913)*	*(0.8327)*	*(0.2838)*	*(0.0013)*
High–Low	Mean	**0.0056****	0.0036	0.0035	0.0008	**0.0062****
	p-value	*(0.0399)*	*(0.1537)*	*(0.1305)*	*(0.7219)*	*(0.0026)*
		4 Factor alpha values				
1	Mean	0.0014	-0.0009	-0.0024	0.0019	-0.0006
	p-value	*(0.6003)*	*(0.6626)*	*(0.2477)*	*(0.3237)*	*(0.7613)*
3	Mean	**0.0071****	0.0022	0.0000	0.0025	**0.0053****
	p-value	*(0.0110)*	*(0.3065)*	*(0.9864)*	*(0.2567)*	*(0.0147)*
High–Low	Mean	**0.0057****	0.0031	0.0024	0.0006	**0.0059****
	p-value	*(0.0454)*	*(0.2291)*	*(0.3103)*	*(0.8066)*	*(0.0057)*
		3F + SUE Factor alpha values				
1	Mean	-0.0005	-0.0025	**-0.0043****	-0.0006	-0.0010
	p-value	*(0.8596)*	*(0.2407)*	*(0.0486)*	*(0.7488)*	*(0.6077)*
3	Mean	0.0033	-0.0004	-0.0017	-0.0002	**0.0036***
	p-value	*(0.3035)*	*(0.8745)*	*(0.4274)*	*(0.9387)*	*(0.0911)*
High–Low	Mean	0.0039	0.0021	0.0026	0.0004	**0.0046****
	p-value	*(0.1748)*	*(0.4151)*	*(0.2823)*	*(0.8549)*	*(0.0311)*
		4F + Liq Factor alpha values				
1	Mean	0.0015	-0.0007	-0.0023	0.0025	0.0002
	p-value	*(0.5908)*	*(0.7344)*	*(0.2796)*	*(0.1944)*	*(0.9087)*
3	Mean	**0.0076****	0.0025	0.0004	0.0030	**0.0057****
	p-value	*(0.0071)*	*(0.2657)*	*(0.8618)*	*(0.1675)*	*(0.0099)*
High–Low	Mean	**0.0061****	0.0032	0.0027	0.0005	**0.0055****
	p-value	*(0.0330)*	*(0.2286)*	*(0.2623)*	*(0.8236)*	*(0.0106)*

Table 2.4 - Raw and risk-adjusted returns for monthly calendar time portfolios triple sorted by SUE, CGO, and abnormal volume

This table reports raw and risk-adjusted returns (Jensen's alphas) from various factor model specifications for high and low abnormal volume tercile portfolios and zero investment portfolios (ZIPs) where portfolios are sorted first by SUE tercile, then by CGO tercile, and finally by abnormal volume tercile. SUE terciles and CGO terciles are calculated across all observations, and abnormal volume terciles are calculated within each CGO tercile. All cutoff values are based on the prior quarter's distribution. The earnings announcement volume premium (EAVP) is defined as the monthly return to a ZIP which takes a long position in abnormal volume tercile 3 (high volume) and a short position in abnormal volume tercile 1 (low volume), within each SUE tercile by CGO tercile. Stocks are equally weighted within each portfolio. For each earnings announcement, we define the earnings announcement window as the three trading day interval, $[t-1, t+1]$, where t is the earnings announcement day. SUE is standardized unexpected earnings based on a seasonal random walk model. The capital gains overhang (CGO) of a stock is defined as the difference between the current stock price and the reference price, which is then normalized by the current stock price as defined in Grinblatt and Han (2005). The reference price is recursively calculated using the prior five years of transaction price and turnover data. We measure the CGO of each company as of five trading days before an earnings announcement. Abnormal volume is defined as the sum of daily market-adjusted volume for the earnings announcement window. Each stock is assigned to a SUE tercile-CGO tercile-abnormal volume tercile portfolio starting from the next month after the end of the earnings announcement window and ending at the end of the next earnings announcement month, or until four months elapse, whichever comes first. Panels A through D report raw and risk-adjusted returns (Jensen's alphas) from various factor model specifications for high and low abnormal volume tercile portfolios and ZIPs. In adjusting for risk factors through regressions, dependent variables for high and low abnormal volume tercile portfolios are raw returns minus the risk-free (t-bill) rate, and dependent variables for ZIPs are raw returns for the high abnormal volume portfolios minus raw returns for the low abnormal volume portfolios. 3F regressions use the standard Fama–French three factors; 4F regressions add the momentum factor as defined in Carhart (1997) to the 3F specification; 4F+Liq Factor regressions add the permanent variable liquidity factor as defined in Sadka (2006) to the 4F regressions. Results with *p-values* below 0.05 (0.10) are marked with ** (*) and are in bold.

	Panel A: Raw returns			*Panel B: 3 Factor alpha values*			*Panel C: 4 Factor alpha values*			*Panel D: 4 Factor + Liq Factor alpha values*		
	SUE Tercile			SUE Tercile			SUE Tercile			SUE Tercile		
CGO Tercile	1	2	3	1	2	3	1	2	3	1	2	3
High volume (abnormal volume tercile 3)												
1	**0.0126****	**0.0134****	**0.0216****	-0.0018	0.0002	**0.0101****	0.0025	0.0033	**0.0145****	0.0025	0.0037	**0.0155****
	(0.0058)	*(0.0007)*	*(0.0000)*	*(0.5161)*	*(0.9448)*	*(0.0009)*	*(0.3526)*	*(0.1500)*	*(0.0000)*	*(0.3565)*	*(0.1149)*	*(0.0000)*
3	**0.0085****	**0.0154****	**0.0191****	-0.0024	**0.0049****	**0.0069****	-0.0023	**0.0037***	**0.0052****	-0.0020	**0.0041****	**0.0056****
	(0.0158)	*(0.0000)*	*(0.0000)*	*(0.2074)*	*(0.0094)*	*(0.0008)*	*(0.2412)*	*(0.0529)*	*(0.0116)*	*(0.3312)*	*(0.0360)*	*(0.0075)*
Low volume (abnormal volume tercile 1)												
1	0.0056	**0.0112****	**0.0128****	**-0.0078****	-0.0017	-0.0006	-0.0025	**0.0031***	0.0048	-0.0026	**0.0035****	0.0050
	(0.1712)	*(0.0029)*	*(0.0049)*	*(0.0020)*	*(0.4138)*	*(0.8469)*	*(0.2339)*	*(0.0726)*	*(0.1078)*	*(0.2267)*	*(0.0449)*	*(0.1014)*
3	**0.0105****	**0.0099****	**0.0139****	-0.0003	-0.0021	0.0016	-0.0012	-0.0022	0.0007	-0.0008	-0.0013	0.0012
	(0.0005)	*(0.0013)*	*(0.0000)*	*(0.8313)*	*(0.2386)*	*(0.3682)*	*(0.4621)*	*(0.2294)*	*(0.6973)*	*(0.6322)*	*(0.4834)*	*(0.5219)*
High volume – low volume												
1	**0.0070****	0.0022	**0.0088****	**0.0060****	0.0019	**0.0107****	**0.0050***	0.0002	**0.0097****	**0.0051***	0.0002	**0.0105****
	(0.0058)	*(0.3783)*	*(0.0117)*	*(0.0218)*	*(0.4578)*	*(0.0031)*	*(0.0627)*	*(0.9361)*	*(0.0089)*	*(0.0617)*	*(0.9482)*	*(0.0051)*
3	-0.0020	**0.0056****	**0.0052****	-0.0021	**0.0070****	**0.0054****	-0.0011	**0.0059****	**0.0045****	-0.0011	**0.0054****	**0.0044****
	(0.3549)	*(0.0136)*	*(0.0071)*	*(0.3393)*	*(0.0012)*	*(0.0037)*	*(0.6211)*	*(0.0071)*	*(0.0165)*	*(0.6122)*	*(0.0159)*	*(0.0203)*

Table 2.5 - Raw and risk-adjusted returns for monthly calendar time portfolios double sorted by size and abnormal volume and triple sorted by size, CGO and abnormal volume

Panel A reports raw and risk-adjusted returns (Jensen's alphas) from various factor model specifications for each abnormal volume tercile portfolio and zero investment portfolios (ZIPs) by size tercile. Size is the market capitalization at the end of the prior calendar year, and size terciles are defined using cutoff values based on the market capitalization distribution of all NYSE companies as of the end of the prior calendar year. For each earnings announcement, we define the earnings announcement window as the three trading day interval, [*t-1*, *t+1*], where *t* is the earnings announcement day. Abnormal volume is defined as the sum of daily market-adjusted volume for the earnings announcement window. Each stock is assigned to an abnormal volume tercile portfolio starting from the next month after the end of the earnings announcement window and ending at the end of the next earnings announcement month, or until four months elapse, whichever comes first. Abnormal volume tercile cutoff values are based on the prior quarter's distribution, and are calculated across all observations. The earnings announcement volume premium (EAVP) is defined as the monthly return to a ZIP which takes a long position in abnormal volume tercile 3 (high volume) and a short position in abnormal volume tercile 1 (low volume). Stocks are equally weighted within each portfolio. In adjusting for risk factors through regressions, dependent variables for abnormal volume tercile portfolios are raw returns minus the risk-free (t-bill) rate, and dependent variables for ZIPs are raw returns for the high abnormal volume portfolios minus raw returns for the low abnormal volume portfolios. 3F regressions use the standard Fama–French three factors; 4F regressions add the momentum factor as defined in Carhart (1997) to the 3F specification; 3F+SUE Factor regressions add a SUE factor to the 3F specification, where the SUE factor is calculated as the difference in monthly equally weighted returns between the highest (decile 10) and the lowest (decile 1) SUE portfolios; 4F+Liq Factor regressions add the permanent variable liquidity factor as defined in Sadka (2006) to the 4F regressions. Panel B reports the same returns where abnormal volume terciles are defined within each capital gains overhang (CGO) tercile, where CGO tercile cutoffs are based on the prior quarter's distribution and are calculated across all observations. Size terciles are defined as in Panel A. The CGO of a stock is defined as the difference between the current stock price and the reference price, which is then normalized by the current stock price as defined in Grinblatt and Han (2005). The reference price is recursively calculated using the prior five years of transaction price and turnover data. We measure the CGO of each company as of five trading days before an earnings announcement. Each stock is assigned to a CGO tercile portfolio starting from the next month after the end of the earnings announcement window and ending at the end of the next earnings announcement month, or until four months elapse, whichever comes first. Results with *p-values* below 0.05 (0.10) are marked with ** (*) and are in bold.

Panel A: Raw and risk adjusted returns by size tercile by abnormal volume tercile

Abnormal Volume Tercile		Size Tercile 1	2	3
		Raw returns		
1	Mean	**0.0095****	**0.0101****	**0.0111****
	p-value	*(0.0079)*	*(0.0026)*	*(0.0002)*
3	Mean	**0.0175****	**0.0143****	**0.0117****
	p-value	*(0.0000)*	*(0.0001)*	*(0.0005)*
High–Low	Mean	**0.0080****	**0.0041****	0.0006
	p-value	*(0.0000)*	*(0.0051)*	*(0.6464)*
		3 Factor alpha values		
1	Mean	**-0.0029***	**-0.0032***	-0.0014
	p-value	*(0.0999)*	*(0.0512)*	*(0.2355)*
3	Mean	**0.0056****	0.0014	-0.0010
	p-value	*(0.0004)*	*(0.3869)*	*(0.4583)*
High–Low	Mean	**0.0085****	**0.0046****	0.0004
	p-value	*(0.0000)*	*(0.0014)*	*(0.7406)*
		4 Factor alpha values		
1	Mean	0.0006	-0.0009	0.0004
	p-value	*(0.6974)*	*(0.5528)*	*(0.7277)*
3	Mean	**0.0070****	0.0023	0.0006
	p-value	*(0.0000)*	*(0.1689)*	*(0.6518)*
High–Low	Mean	**0.0065****	**0.0032****	0.0002
	p-value	*(0.0001)*	*(0.0248)*	*(0.8686)*
		3F + SUE Factor alpha values		
1	Mean	-0.0020	**-0.0030***	-0.0014
	p-value	*(0.2748)*	*(0.0874)*	*(0.2627)*
3	Mean	**0.0045****	0.0010	-0.0018
	p-value	*(0.0061)*	*(0.5426)*	*(0.1979)*
High–Low	Mean	**0.0065****	**0.0040****	-0.0004
	p-value	*(0.0001)*	*(0.0079)*	*(0.7402)*
		4F + Liq Factor alpha values		
1	Mean	0.0009	-0.0009	0.0006
	p-value	*(0.5671)*	*(0.5723)*	*(0.5863)*
3	Mean	**0.0073****	**0.0029***	0.0008
	p-value	*(0.0000)*	*(0.0833)*	*(0.5614)*
High–Low	Mean	**0.0064****	**0.0038****	0.0002
	p-value	*(0.0001)*	*(0.0084)*	*(0.8976)*

Panel B: Raw and risk-adjusted returns for portfolios sorted by size by CGO by abnormal volume

Abnormal Volume Tercile		Size Tercile 1			Size Tercile 2			Size Tercile 3		
		CGO Tercile			CGO Tercile			CGO Tercile		
		1	2	3	1	2	3	1	2	3
		Raw returns			*Raw returns*			*Raw returns*		
1	Mean	0.0086**	0.0082**	0.0118**	0.0087**	0.0098**	0.0111**	0.0108**	0.0097**	0.0131**
	p-value	*(0.0410)*	*(0.0206)*	*(0.0004)*	*(0.0457)*	*(0.0033)*	*(0.0006)*	*(0.0073)*	*(0.0011)*	*(0.0000)*
3	Mean	0.0180**	0.0136**	0.0190**	0.0105**	0.0129**	0.0172**	0.0131**	0.0104**	0.0117**
	p-value	*(0.0000)*	*(0.0005)*	*(0.0000)*	*(0.0227)*	*(0.0006)*	*(0.0000)*	*(0.0019)*	*(0.0022)*	*(0.0004)*
High–Low	Mean	0.0093**	0.0054**	0.0072**	0.0019	0.0031	0.0061**	0.0022	0.0007	-0.0014
	p-value	*(0.0001)*	*(0.0414)*	*(0.0069)*	*(0.5757)*	*(0.1339)*	*(0.0198)*	*(0.3798)*	*(0.6368)*	*(0.4620)*
		3 Factor alpha values			*3 Factor alpha values*			*3 Factor alpha values*		
1	Mean	-0.0046*	-0.0036	0.0007	-0.0055*	-0.0029	-0.0007	-0.0021	-0.0026*	0.0012
	p-value	*(0.0578)*	*(0.1064)*	*(0.7593)*	*(0.0525)*	*(0.1579)*	*(0.7091)*	*(0.4797)*	*(0.0743)*	*(0.4845)*
3	Mean	0.0055**	0.0024	0.0076**	-0.0033	-0.0003	0.0054**	-0.0010	-0.0025	0.0003
	p-value	*(0.0274)*	*(0.2599)*	*(0.0008)*	*(0.3022)*	*(0.8686)*	*(0.0259)*	*(0.6900)*	*(0.1045)*	*(0.8588)*
High–Low	Mean	0.0101**	0.0060**	0.0069**	0.0022	0.0026	0.0062**	0.0011	0.0001	-0.0009
	p-value	*(0.0001)*	*(0.0252)*	*(0.0078)*	*(0.5185)*	*(0.2075)*	*(0.0191)*	*(0.6657)*	*(0.9360)*	*(0.6389)*
		4 Factor alpha values			*4 Factor alpha values*			*4 Factor alpha values*		
1	Mean	0.0002	-0.0014	0.0009	-0.0005	-0.0016	-0.0013	0.0045*	-0.0014	-0.0002
	p-value	*(0.9072)*	*(0.5310)*	*(0.6954)*	*(0.8513)*	*(0.4480)*	*(0.5105)*	*(0.0672)*	*(0.3514)*	*(0.9122)*
3	Mean	0.0089**	0.0031	0.0065**	0.0009	-0.0004	0.0034	0.0044**	-0.0017	-0.0010
	p-value	*(0.0002)*	*(0.1659)*	*(0.0045)*	*(0.7747)*	*(0.8626)*	*(0.1616)*	*(0.0313)*	*(0.2867)*	*(0.5495)*
High–Low	Mean	0.0087**	0.0044	0.0056**	0.0014	0.0012	0.0047*	0.0000	-0.0003	-0.0008
	p-value	*(0.0006)*	*(0.1028)*	*(0.0333)*	*(0.6999)*	*(0.5548)*	*(0.0771)*	*(0.9922)*	*(0.8402)*	*(0.6796)*
		3F + SUE Factor alpha values			*3F + SUE Factor alpha values*			*3F + SUE Factor alpha values*		
1	Mean	-0.0024	-0.0028	-0.0019	-0.0032	-0.0036*	-0.0027	0.0004	-0.0033**	-0.0002
	p-value	*(0.3455)*	*(0.2196)*	*(0.3867)*	*(0.2823)*	*(0.0998)*	*(0.1718)*	*(0.8913)*	*(0.0342)*	*(0.8855)*
3	Mean	0.0061**	0.0020	0.0040*	-0.0001	-0.0013	0.0029	0.0005	-0.0038**	-0.0024
	p-value	*(0.0200)*	*(0.3853)*	*(0.0695)*	*(0.9767)*	*(0.5317)*	*(0.2480)*	*(0.8362)*	*(0.0194)*	*(0.1396)*
High–Low	Mean	0.0084**	0.0048*	0.0060**	0.0031	0.0022	0.0056**	0.0001	-0.0005	-0.0021
	p-value	*(0.0011)*	*(0.0858)*	*(0.0282)*	*(0.3957)*	*(0.2987)*	*(0.0426)*	*(0.9654)*	*(0.7508)*	*(0.2746)*
		4F + Liq Factor alpha values			*4F + Liq Factor alpha values*			*4F + Liq Factor alpha values*		
1	Mean	0.0004	-0.0009	0.0012	-0.0006	-0.0015	-0.0008	0.0048*	-0.0012	0.0006
	p-value	*(0.8642)*	*(0.6879)*	*(0.5914)*	*(0.8091)*	*(0.4682)*	*(0.6725)*	*(0.0520)*	*(0.4159)*	*(0.7272)*
3	Mean	0.0093**	0.0033	0.0068**	0.0012	0.0004	0.0040	0.0046**	-0.0015	-0.0007
	p-value	*(0.0001)*	*(0.1429)*	*(0.0036)*	*(0.6975)*	*(0.8574)*	*(0.1074)*	*(0.0273)*	*(0.3623)*	*(0.6888)*
High–Low	Mean	0.0090**	0.0042	0.0056**	0.0019	0.0019	0.0048*	-0.0002	-0.0002	-0.0012
	p-value	*(0.0005)*	*(0.1299)*	*(0.0389)*	*(0.6069)*	*(0.3563)*	*(0.0755)*	*(0.9434)*	*(0.8767)*	*(0.5223)*

Table 2.6 - Raw and risk-adjusted returns for monthly calendar time portfolios double sorted by CGO and abnormal volume (December and January removed)

This table reports raw and risk-adjusted returns (Jensen's alphas) from various factor model specifications for high and low abnormal volume tercile portfolios and zero investment portfolios (ZIPs) by CGO quintile after removing December and January observations from each portfolio. The earnings announcement volume premium (EAVP) is defined as the monthly return to a ZIP which takes a long position in abnormal volume tercile 3 (high volume) and a short position in abnormal volume tercile 1 (low volume), where abnormal volume terciles are defined within each CGO quintile. Stocks are equally weighted within each portfolio. For each earnings announcement, we define the earnings announcement window as the three trading day interval, $[t\text{-}1, t\text{+}1]$, where t is the earnings announcement day. The capital gains overhang (CGO) of a stock is defined as the difference between the current stock price and the reference price, which is then normalized by the current stock price as defined in Grinblatt and Han (2005). The reference price is recursively calculated using the prior five years of transaction price and turnover data. We measure the CGO of each company as of five trading days before an earnings announcement. Each stock is assigned to a CGO quintile portfolio starting from the next month after the end of the earnings announcement window and ending at the end of the next earnings announcement month, or until four months elapse, whichever comes first. Abnormal volume is defined as the sum of daily market-adjusted volume for the earnings announcement window. Within each CGO quintile, abnormal volume terciles are assigned in the same manner as CGO quintiles. All cutoff values are based on the prior quarter's distribution. In adjusting for risk factors through regressions, dependent variables for high and low abnormal volume tercile portfolios are raw returns minus the risk-free (t-bill) rate, and dependent variables for ZIPs are raw returns for the high abnormal volume portfolios minus raw returns for the low abnormal volume portfolios. 3F regressions use the standard Fama–French three factors; 4F regressions add the momentum factor as defined in Carhart (1997) to the 3F specification; 3F+SUE Factor regressions add a SUE factor to the 3F specification, where the SUE factor is calculated as the difference in monthly equally weighted returns between the highest (decile 10) and the lowest (decile 1) SUE portfolios; 4F+Liq Factor regressions add the permanent variable liquidity factor as defined in Sadka (2006) to the 4F regressions. Results with *p-values* below 0.05 (0.10) are marked with ** (*) and are in bold.

Abnormal Volume Tercile		CGO Quintile 1	2	3	4	5
		Raw returns				
1	Mean	0.0040	**0.0064***	0.0050	**0.0091****	**0.0102****
	p-value	*(0.3753)*	*(0.0647)*	*(0.1232)*	*(0.0048)*	*(0.0016)*
3	Mean	**0.0115****	**0.0091****	**0.0077****	**0.0099****	**0.0146****
	p-value	*(0.0119)*	*(0.0289)*	*(0.0388)*	*(0.0076)*	*(0.0001)*
High–Low	Mean	**0.0075****	0.0026	0.0027	0.0008	**0.0044****
	p-value	*(0.0010)*	*(0.2120)*	*(0.1343)*	*(0.6409)*	*(0.0157)*
		3 Factor alpha values				
1	Alpha	**-0.0054****	**-0.0030***	**-0.0042****	-0.0004	0.0020
	p-value	*(0.0261)*	*(0.0966)*	*(0.0146)*	*(0.7635)*	*(0.1914)*
3	Alpha	0.0021	0.0000	-0.0010	0.0009	**0.0072****
	p-value	*(0.4501)*	*(0.9988)*	*(0.5354)*	*(0.5988)*	*(0.0000)*
High–Low	Alpha	**0.0074****	0.0030	**0.0032***	0.0013	**0.0052****
	p-value	*(0.0014)*	*(0.1265)*	*(0.0605)*	*(0.4160)*	*(0.0022)*
		4 Factor alpha values				
1	Alpha	0.0001	0.0012	-0.0012	0.0002	0.0014
	p-value	*(0.9783)*	*(0.4684)*	*(0.4476)*	*(0.9166)*	*(0.3693)*
3	Alpha	**0.0080****	0.0021	0.0005	0.0021	**0.0065****
	p-value	*(0.0015)*	*(0.2897)*	*(0.7808)*	*(0.2234)*	*(0.0002)*
High–Low	Alpha	**0.0079****	0.0009	0.0017	0.0019	**0.0050****
	p-value	*(0.0013)*	*(0.6524)*	*(0.3279)*	*(0.2549)*	*(0.0052)*
		3F + SUE Factor alpha values				
1	Alpha	-0.0030	-0.0016	**-0.0043****	**-0.0029***	0.0010
	p-value	*(0.2390)*	*(0.4026)*	*(0.0170)*	*(0.0505)*	*(0.5490)*
3	Alpha	0.0044	-0.0007	-0.0018	-0.0017	**0.0046****
	p-value	*(0.1264)*	*(0.7505)*	*(0.2819)*	*(0.3064)*	*(0.0069)*
High–Low	Alpha	**0.0074****	0.0010	0.0025	0.0012	**0.0036****
	p-value	*(0.0030)*	*(0.6439)*	*(0.1625)*	*(0.4882)*	*(0.0436)*
		4F + Liq Factor alpha values				
1	Alpha	0.0001	0.0012	-0.0013	0.0007	0.0016
	p-value	*(0.9793)*	*(0.4497)*	*(0.4146)*	*(0.6459)*	*(0.3237)*
3	Alpha	**0.0082****	0.0021	0.0005	0.0024	**0.0066****
	p-value	*(0.0011)*	*(0.2831)*	*(0.7496)*	*(0.1758)*	*(0.0002)*
High–Low	Alpha	**0.0082****	0.0009	0.0019	0.0016	**0.0050****
	p-value	*(0.0009)*	*(0.6599)*	*(0.2839)*	*(0.3354)*	*(0.0061)*

Table 2.7 - Raw and risk-adjusted returns for monthly calendar time portfolios double sorted by CGO and abnormal volume (subsample analyses by decade)

Panel A and Panel B report raw and risk-adjusted returns (Jensen's alphas) from various factor model specifications for high and low abnormal volume tercile portfolios and zero investment portfolios (ZIPs) by CGO quintile for the subsample periods of 1983–1990 and 1991–2000, respectively. The earnings announcement volume premium (EAVP) is defined as the monthly return to a ZIP which takes a long position in abnormal volume tercile 3 (high volume) and a short position in abnormal volume tercile 1 (low volume), where abnormal volume terciles are defined within each CGO quintile. Stocks are equally weighted within each portfolio. For each earnings announcement, we define the earnings announcement window as the three trading day interval, [*t-1*, *t+1*], where *t* is the earnings announcement day. The capital gains overhang (CGO) of a stock is defined as the difference between the current stock price and the reference price, which is then normalized by the current stock price as defined in Grinblatt and Han (2005). The reference price is recursively calculated using the prior five years of transaction price and turnover data. We measure the CGO of each company as of five trading days before an earnings announcement. Each stock is assigned to a CGO quintile portfolio starting from the next month after the end of the earnings announcement window and ending at the end of the next earnings announcement month, or until four months elapse, whichever comes first. Abnormal volume is defined as the sum of daily market-adjusted volume for the earnings announcement window. Within each CGO quintile, abnormal volume terciles are assigned in the same manner as CGO quintiles. All cutoff values are based on the prior quarter's distribution. In adjusting for risk factors through regressions, dependent variables for high and low abnormal volume tercile portfolios are raw returns minus the risk-free (t-bill) rate, and dependent variables for ZIPs are raw returns for the high abnormal volume portfolios minus raw returns for the low abnormal volume portfolios. 3F regressions use the standard Fama–French three factors; 4F regressions add the momentum factor as defined in Carhart (1997) to the 3F specification; 3F+SUE Factor regressions add a SUE factor to the 3F specification, where the SUE factor is calculated as the difference in monthly equally weighted returns between the highest (decile 10) and the lowest (decile 1) SUE portfolios; 4F+Liq Factor regressions add the permanent variable liquidity factor as defined in Sadka (2006) to the 4F regressions. Results with *p-values* below 0.05 (0.10) are marked with ** (*) and are in bold.

Panel A: 1983 - 1990 only

Abnormal Volume Tercile		CGO Quintile 1	2	3	4	5
		Raw returns				
1	Mean	0.0044	**0.0096***	0.0087	**0.0137****	**0.0106***
	p-value	*(0.5159)*	*(0.0614)*	*(0.1087)*	*(0.0163)*	*(0.0521)*
3	Mean	0.0104	0.0089	0.0094	**0.0118****	**0.0150****
	p-value	*(0.1217)*	*(0.1432)*	*(0.1194)*	*(0.0495)*	*(0.0109)*
High–Low	Mean	**0.0060****	-0.0007	0.0007	-0.0019	**0.0044****
	p-value	*(0.0285)*	*(0.7641)*	*(0.7627)*	*(0.4711)*	*(0.0305)*
		3 Factor alpha values				
1	Alpha	-0.0032	0.0011	-0.0001	**0.0042****	**0.0038***
	p-value	*(0.2397)*	*(0.5210)*	*(0.9500)*	*(0.0226)*	*(0.0632)*
3	Alpha	0.0035	-0.0006	0.0012	**0.0039***	**0.0083****
	p-value	*(0.2626)*	*(0.7559)*	*(0.5285)*	*(0.0665)*	*(0.0000)*
High–Low	Alpha	**0.0067****	-0.0017	0.0013	-0.0003	**0.0046****
	p-value	*(0.0200)*	*(0.4692)*	*(0.5770)*	*(0.9110)*	*(0.0286)*
		4 Factor alpha values				
1	Alpha	-0.0006	0.0022	0.0002	**0.0037****	0.0019
	p-value	*(0.7934)*	*(0.1785)*	*(0.8874)*	*(0.0436)*	*(0.2685)*
3	Alpha	**0.0058****	0.0001	0.0001	0.0025	**0.0076****
	p-value	*(0.0445)*	*(0.9659)*	*(0.9544)*	*(0.2005)*	*(0.0000)*
High–Low	Alpha	**0.0064****	-0.0022	-0.0001	-0.0012	**0.0057****
	p-value	*(0.0282)*	*(0.3630)*	*(0.9510)*	*(0.6588)*	*(0.0049)*
		3F + SUE Factor alpha values				
1	Alpha	0.0024	**0.0033***	-0.0002	0.0018	0.0016
	p-value	*(0.3512)*	*(0.0716)*	*(0.9031)*	*(0.3439)*	*(0.4484)*
3	Alpha	**0.0068****	0.0006	-0.0004	0.0007	**0.0075****
	p-value	*(0.0428)*	*(0.7614)*	*(0.8449)*	*(0.7455)*	*(0.0002)*
High–Low	Alpha	0.0044	-0.0027	-0.0002	-0.0011	**0.0059****
	p-value	*(0.1476)*	*(0.2860)*	*(0.9448)*	*(0.7104)*	*(0.0098)*
		4F + Liq Factor alpha values				
1	Alpha	-0.0009	0.0024	0.0001	**0.0044****	0.0019
	p-value	*(0.7083)*	*(0.1599)*	*(0.9657)*	*(0.0194)*	*(0.2826)*
3	Alpha	**0.0059****	0.0000	0.0002	0.0029	**0.0076****
	p-value	*(0.0461)*	*(0.9879)*	*(0.9343)*	*(0.1505)*	*(0.0001)*
High–Low	Alpha	**0.0068****	-0.0024	0.0001	-0.0014	**0.0058****
	p-value	*(0.0234)*	*(0.3189)*	*(0.9722)*	*(0.5994)*	*(0.0054)*

Panel B: 1991 - 2000 only

Abnormal Volume Tercile		CGO Quintile 1	2	3	4	5
		Raw returns				
1	Mean	**0.0120****	**0.0105****	**0.0083****	**0.0117****	**0.0137****
	p-value	*(0.0204)*	*(0.0073)*	*(0.0094)*	*(0.0005)*	*(0.0001)*
3	Mean	**0.0226****	**0.0167****	**0.0138****	**0.0155****	**0.0195****
	p-value	*(0.0000)*	*(0.0003)*	*(0.0003)*	*(0.0002)*	*(0.0000)*
High–Low	Mean	**0.0106****	**0.0063****	**0.0055****	0.0038	**0.0058****
	p-value	*(0.0007)*	*(0.0240)*	*(0.0226)*	*(0.1070)*	*(0.0243)*
		3 Factor alpha values				
1	Alpha	-0.0059	**-0.0059****	**-0.0055****	-0.0029	-0.0001
	p-value	*(0.1006)*	*(0.0150)*	*(0.0099)*	*(0.1289)*	*(0.9663)*
3	Alpha	0.0045	0.0000	-0.0011	-0.0008	**0.0048***
	p-value	*(0.2183)*	*(0.9968)*	*(0.6037)*	*(0.7423)*	*(0.0607)*
High–Low	Alpha	**0.0104****	**0.0059****	**0.0044***	0.0021	**0.0048****
	p-value	*(0.0018)*	*(0.0297)*	*(0.0536)*	*(0.3481)*	*(0.0377)*
		4 Factor alpha values				
1	Alpha	0.0030	-0.0009	-0.0025	-0.0018	0.0005
	p-value	*(0.2857)*	*(0.6810)*	*(0.2341)*	*(0.3702)*	*(0.8213)*
3	Alpha	**0.0124****	0.0028	0.0008	0.0009	0.0044
	p-value	*(0.0001)*	*(0.3065)*	*(0.7201)*	*(0.7021)*	*(0.1040)*
High–Low	Alpha	**0.0095****	0.0037	0.0033	0.0027	0.0039
	p-value	*(0.0070)*	*(0.1847)*	*(0.1725)*	*(0.2504)*	*(0.1099)*
		3F + SUE Factor alpha values				
1	Alpha	-0.0048	**-0.0053****	**-0.0060****	**-0.0047****	-0.0015
	p-value	*(0.1964)*	*(0.0373)*	*(0.0073)*	*(0.0125)*	*(0.4411)*
3	Alpha	0.0061	-0.0009	-0.0023	-0.0032	0.0026
	p-value	*(0.1122)*	*(0.7452)*	*(0.3052)*	*(0.1819)*	*(0.3032)*
High–Low	Alpha	**0.0109****	0.0043	0.0037	0.0016	**0.0041***
	p-value	*(0.0018)*	*(0.1187)*	*(0.1172)*	*(0.4965)*	*(0.0895)*
		4F + Liq Factor alpha values				
1	Alpha	0.0029	-0.0009	-0.0025	-0.0018	0.0005
	p-value	*(0.2903)*	*(0.6767)*	*(0.2290)*	*(0.3739)*	*(0.8197)*
3	Alpha	**0.0124****	0.0028	0.0008	0.0009	0.0044
	p-value	*(0.0001)*	*(0.3090)*	*(0.7270)*	*(0.7085)*	*(0.1048)*
High–Low	Alpha	**0.0095****	0.0037	0.0033	0.0027	0.0039
	p-value	*(0.0073)*	*(0.1849)*	*(0.1735)*	*(0.2550)*	*(0.1113)*

Table 2.8 - Drift regressions

This table reports the results of regressing sixty day cumulative abnormal returns (DRIFT60) on various explanatory variables, using weighted least squares (WLS) regressions with weights equal to (1/number of observations in each quarter). DRIFT60 is the sum of daily abnormal returns from t+2 to t+61, where t is the earnings announcement day. Daily abnormal returns are calculated using size- and B/M-matched portfolios as defined in Brav and Gompers (1997). SUE is standardized unexpected earnings based on a seasonal random walk model. Size is the market capitalization at the end of the prior calendar year (in millions of dollars) and LSIZE is the natural logarithm of (1+size). LSIZE_SUE is the interaction term of LSIZE and SUE. For each earnings announcement, we define the earnings announcement window as the three trading day interval, [t-1, t+1], where t is the earnings announcement day. Abnormal volume is the sum of daily market-adjusted volume for the earnings announcement window, and VOL is the abnormal volume decile number defined in each quarter using cutoff values based on the prior quarter's distribution. CGO is the difference between the current stock price and the reference price, which is then normalized by the current stock price as defined in Grinblatt and Han (2005). The reference price is recursively calculated using the prior five years of transaction price and turnover data. We measure the CGO of each company as of five trading days before an earnings announcement. CGOPOS is max(0,CGO) and CGONEG is -min(0,CGO). VOL_CGO, VOL_CGOPOS, and VOL_CGONEG are VOL interacted with CGO, CGOPOS, and CGONEG, respectively. Turnover is the average daily share turnover over the 52-week period ending immediately prior to the week in which the earnings announcement is made, and LTO is the natural logarithm of (1+turnover). Momentum (MOM) is compounded returns (inclusive of dividends) for the 12 calendar months ending immediately prior to the month in which the earnings announcement is made. MOMSQ is MOM squared, and VOL_MOMSQ is the interaction term of abnormal volume and MOMSQ. Regressions include year-quarter dummies, and standard errors are calculated by clustering observations by year-quarter. The table reports parameter estimates (top) and *p-values* (bottom, in italics) for each regression. Results with *p-values* below 0.05 (0.10) are marked with ** (*) and are in bold.

Regression Number	SUE	LSIZE_SUE	VOL	VOL_ CGO	VOL_ CGOPOS	VOL_ CGONEG	LTO	MOM	VOL_ MOMSQ	ADJ. R-SQ.
1	**0.0147****	**-0.0017****	**0.0022****							0.0096
	(0.0000)	*(0.0000)*	*(0.0000)*							
2	**0.0162****	**-0.0019****	**0.0024****	**-0.0038****						0.0129
	(0.0000)	*(0.0000)*	*(0.0000)*	*(0.0000)*						
3	**0.0160****	**-0.0018****	**0.0009****		**0.0047****	**0.0075****				0.0160
	(0.0000)	*(0.0000)*	*(0.0320)*		*(0.0030)*	*(0.0000)*				
4	**0.0157****	**-0.0018****	**0.0009****		**0.0047****	**0.0080****	0.0039	0.0069	-0.0006	0.0162
	(0.0000)	*(0.0000)*	*(0.0210)*		*(0.0070)*	*(0.0000)*	*(0.5230)*	*(0.2480)*	*(0.2450)*	

Figure 2.1 - Daily market-adjusted volume around event time

This figure shows the average of daily market-adjusted volume by trading day relative to an earnings announcement day (*t*) for the full sample, good news and bad news subsamples (top and bottom SUE quintiles for each quarter, respectively) and high and low CGO subsamples (top and bottom CGO quintiles for each quarter, respectively). SUE is standardized unexpected earnings based on a seasonal random walk model. The capital gains overhang (CGO) of a stock is defined as the difference between the current stock price and the reference price, which is then normalized by the current stock price as defined in Grinblatt and Han (2005). The reference price is recursively calculated using the prior five years of transaction price and turnover data. We measure the CGO of each company as of five trading days before an earnings announcement. All cutoff values are based on the prior quarter's distribution.

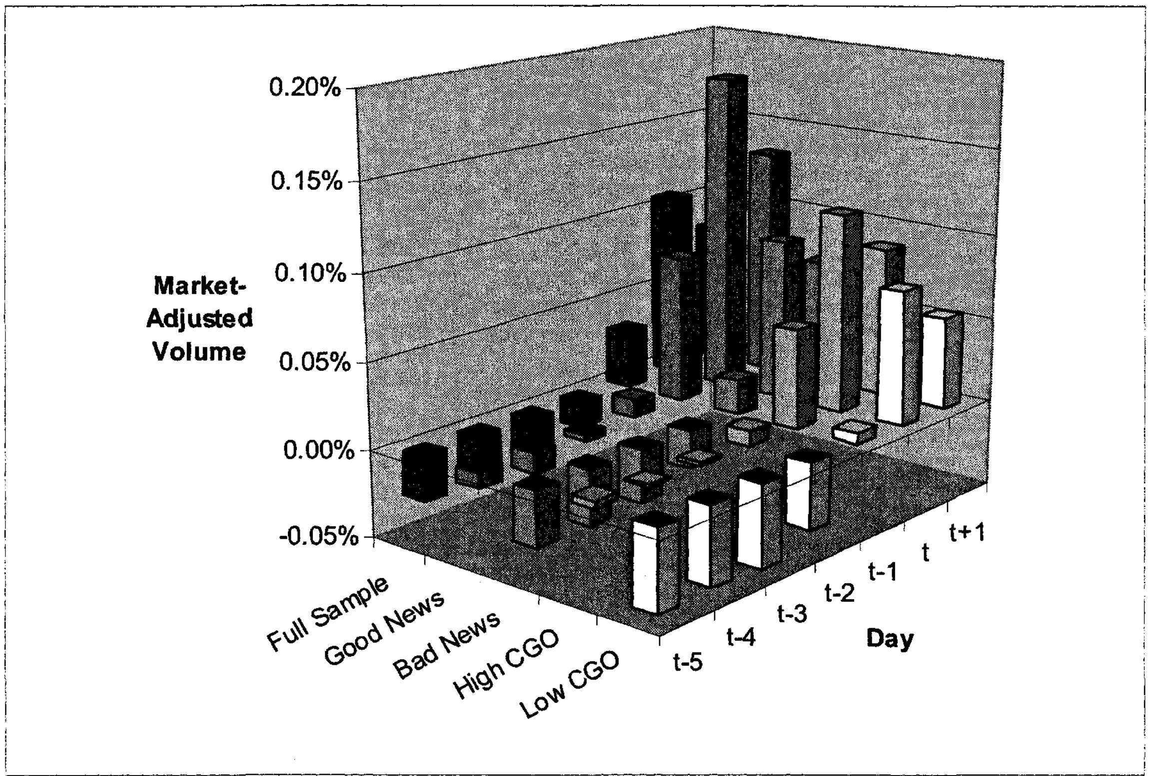


Figure 2.2 - Timeline for variable construction

This figure shows the trading day ranges used in calculating the main portfolio sorting variables when an earnings announcement occurs on day *t*. For each earnings announcement, we define the earnings announcement window as the three trading day interval, [*t-1*, *t+1*]. Abnormal volume is defined as the sum of daily market-adjusted volume for the earnings announcement window. The capital gains overhang (CGO) of a stock is defined as the difference between the current stock price and the reference price, which is then normalized by the current stock price as defined in Grinblatt and Han (2005). The reference price is recursively calculated using the prior five years of transaction price and turnover data. We measure the CGO of each company as of five trading days before an earnings announcement.

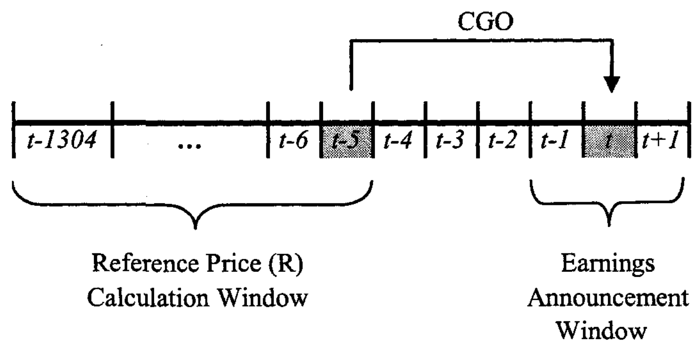
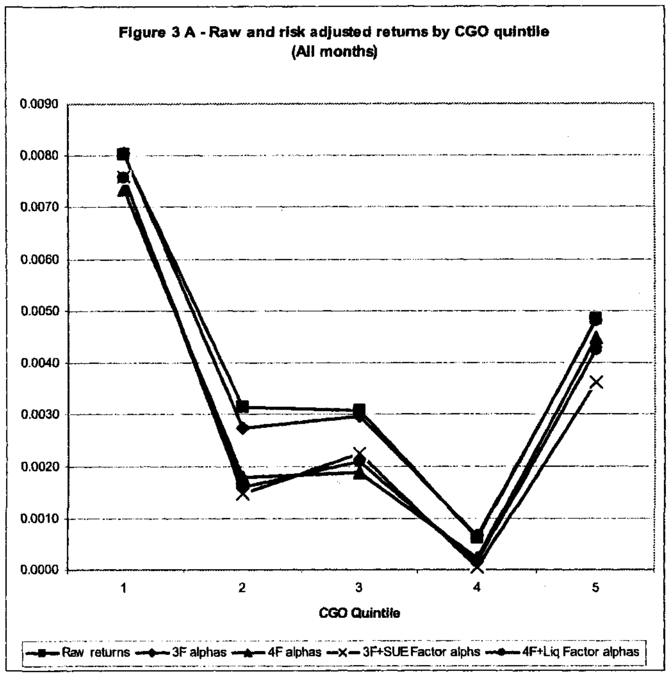
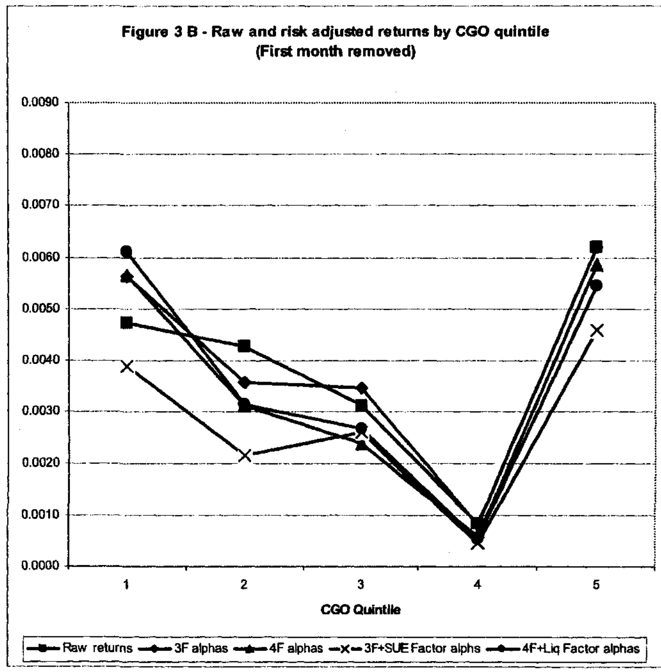


Figure 2.3 - Raw and risk-adjusted returns for monthly calendar time portfolios double sorted by CGO and abnormal volume

This figure shows the magnitudes of the earnings announcement volume premium (EAVP) by CGO quintile from various factor model specifications. The earnings announcement volume premium (EAVP) is defined as the monthly return to a zero investment portfolio (ZIP) which takes a long position in abnormal volume tercile 3 (high volume) and a short position in abnormal volume tercile 1 (low volume), where abnormal volume terciles are defined within each CGO quintile. Stocks are equally weighted within each portfolio. For each earnings announcement, we define the earnings announcement window as the three trading day interval, [*t-1*, *t+1*], where *t* is the earnings announcement day. The capital gains overhang (CGO) of a stock is defined as the difference between the current stock price and the reference price, which is then normalized by the current stock price as defined in Grinblatt and Han (2005). The reference price is recursively calculated using the prior five years of transaction price and turnover data. We measure the CGO of each company as of five trading days before an earnings announcement. Abnormal volume is defined as the sum of daily market-adjusted volume for the earnings announcement window. Figure 3 A shows the EAVP values under the original portfolio strategy, where each stock is assigned to a CGO quintile-abnormal volume tercile portfolio starting from the next month after the end of the earnings announcement window and ending at the end of the next earnings announcement month, or until four months elapse, whichever comes first. Figure 3 B shows the EAVP values after removing the first month of the holding period return for each stock in each portfolio. All cutoff values are based on the prior quarter's distribution. In adjusting for risk factors through regressions, dependent variables for high and low abnormal volume tercile portfolios are raw returns minus the risk-free (t-bill) rate, and dependent variables for ZIPs are raw returns for the high abnormal volume portfolios minus raw returns for the low abnormal volume portfolios. 3F regressions use the standard Fama–French three factors; 4F regressions add the momentum factor as defined in Carhart (1997) to the 3F specification; 3F+SUE Factor regressions add a SUE factor to the 3F specification, where the SUE factor is calculated as the difference in monthly equally weighted returns between the highest (decile 10) and the lowest (decile 1) SUE portfolios; 4F+Liq Factor regressions add the permanent variable liquidity factor as defined in Sadka (2006) to the 4F regressions.





BIBLIOGRAPHY

Ajinkya, Bipin B., Rowland K. Atiase, Michael J. Gift, 1991, Volume of Trading and the Dispersion in Financial Analysts' Earnings Forecasts, *Accounting Review* 66, 389-401

Amihud, Yakov, 2002, Illiquidity and Stock Returns: Cross-section and Time-series Effects, *Journal of Financial Markets* 5, 31-56.

Avramov, Doron, Tarun Chordia, Gergana Jostova, and Alexander Philipov, 2008, Dispersion in Analysts' Earnings Forecasts and Credit Rating, forthcoming in *Journal of Financial Economics*.

Baker, Malcolm, and Jeremy Stein, 2004, Market Liquidity as a Sentiment Indicator, *Journal of Financial Markets* 7, 271-299.

Bamber, Linda S., and Youngsoon S. Cheon, 1995, Differential Price and Volume Reactions to Accounting Earnings Announcements, *Accounting Review* 70, 417-441.

Barber, Brad M., Yi-Tsung Lee, Yu-Jane Liu, and Terrance Odean, 2007, Is the Aggregate Investor Reluctant to Realize Losses? Evidence from Taiwan, *European Financial Management* 13, 423-447.

Barber, Brad M. and Terrance Odean, 2008, All that Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors, *Review of Financial Studies* 21, 763-784.

Barberis, Nicholas, and Ming Huang, 2001, Mental Accounting, Loss Aversion and Individual Stock Returns, *Journal of Finance* 56, 1247-1292.

Barberis, Nicholas, Huang, Ming and Tano Santos 2001, Prospect Theory and Asset Prices, *Quarterly Journal of Economics* 116, 1-53.

Bernard, Victor and Jacob Thomas, 1989, Post-earnings Announcement Drift: Delayed Price Response or Risk Premium?, *Journal of Accounting Research* 27, 1-36.

Bernard, Victor and Jacob Thomas, 1990, Evidence that Stock Prices Do Not Fully Reflect the Implications of Current Earnings for Future Earnings, *Journal of Accounting and Economics* 13, 305-340.

Brav, Alon, and Paul A. Gompers, 1997, Myth or Reality? The Long-run Underperformance of Initial Public Offerings: Evidence from Venture and Nonventure Capital-backed Companies, *Journal of Finance* 52, 1791-1821.

Brav, Alon and J.B. Heaton, 2002, Competing Theories of Financial Anomalies, *Review of Financial Studies* 15, 575-606.

Campbell, John Y., Sanford J. Grossman, and Jiang Wang, 1993, Trading Volume and Serial Correlation in Stock Returns, *Quarterly Journal of Economics* 108, 905-939.

Carhart, Mark M., 1997, On Persistence in Mutual Fund Performance, *Journal of Finance* 52, 57-82.

Chordia, Tarun and Lakshmanan Shivakumar, 2006, Earnings and Price Momentum, *Journal of Financial Economics* 80, 627-656.

Chordia, Tarun, Amit Goyal, Gil Sadka, Ronnie Sadka, and Lakshmanan Shivakumar, 2007a, Liquidity and the Post-Earnings-Announcement Drift, Emory University working paper.

Chordia, Tarun, Sahn-Wook Huh, and Avanidhar Subrahmanyam, 2007b, The Cross-section of Expected Trading Activity, *Review of Financial Studies* 20, 709-740.

Ciccone, Stephen J., 2001, Analyst Forecast Properties, Financial Distress, and Business Risk, University of New Hampshire working paper.

Daniel, Kent D., David Hirshleifer, and Avanidhar Subrahmanyam, 1998, A Theory of Investor Overconfidence, Self-Attribution, and Security Returns Anomalies, *Journal of Finance* 53, 1839-1885.

Diether, Karl, Christopher Malloy, and Anna Scherbina, 2002, Differences of Opinion and the Cross Section of Stock Returns, *Journal of Finance* 57, 2113-2141.

Fama, Eugene, Kenneth R. French, 1993, Common Risk Factors in the Returns on Stocks and Bonds, *Journal of Financial Economics* 33, 3-56.

Francis, Jennifer, Ryan Lafond, Per Olsson and Katherine Schipper, 2007, Information Uncertainty and Post-earnings-announcement-drift, *Journal of Business Finance & Accounting* 34, 403-433.

Frazzini, Andrea, 2006, The Disposition Effect and Underreaction to News, *Journal of Finance* 61, 2017-2046.

Lamont, Owen, and Andrea Frazzini, 2007, The Earnings Announcement Premium and Trading Volume, NBER working paper w13090.

Garfinkel, Jon, and Jonathan Sokobin, 2006, Volume, opinion divergence, and returns: A study of post-earnings announcement drift, Journal of Accounting Research 44, 85-112.

Gervais, Simon, Ron Kaniel, and Dan Mingelgrin, 2001, The High-volume Return Premium, *Journal of Finance* 56, 877-919.

Gomes, Francisco J., 2005, Portfolio Choice and Trading Volume with Loss-averse Investors, *Journal of Business* 78, 675-706.

Grinblatt, Mark, and Bing Han, 2005, Prospect Theory, Mental Accounting, and Momentum, *Journal of Financial Economics* 78, 311-339.

Grinblatt, Mark, and Matti Keloharju, 2001, What Makes Investors Trade?, *Journal of Finance* 56, 585-616.

Grinblatt, Mark, and Matti Keloharju, 2004, Tax-loss Trading and Wash Sales, *Journal of Financial Economics* 71, 51-76.

Grinblatt, Mark and Tobias J. Moskowitz, 2004, Predicting Stock Price Movements from Past Returns: The Role of Consistency and Tax-loss Selling, *Journal of Financial Economics* 71, 541-579.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: implications for stock market efficiency, Journal of Finance 48, 65-91.

Jin, Li, 2006, Capital Gain Tax Overhang and Price Pressure, *Journal of Finance* 61, 1399-1430.

Kandel, Eugene, and Neil D. Pearson, 1995, Differential Interpretation of Public Signals and Trade in Speculative Markets, *Journal of Political Economy* 103, 831-872.

Kaniel, Ron, Dong Li, and Laura Starks, 2005, Investor Visibility Events: Cross-country Evidence, University of Texas working paper.

Karpoff, Jonathan M., 1987, The Relation Between Price Changes and Trading Volume: A Survey, *Journal of Financial and Quantitative Analysis* 22, 109-126.

Lee, Charles M., Belinda Mucklow, and Mark Ready, 1993, Spreads, Depths, and the Impact of Earnings Information: An Intraday Analysis, *Review of Financial Studies* 6, 345-374.

Lerman, Alina, Joshua Livnat, and Richard R. Mendenhall, 2008, The High-Volume Return Premium and Post-Earnings Announcement Drift, Stern School of Business Administration, New York University working paper.

Llorente, Guillermo, Roni Michaely, Gideon Saar, and Jiang Wang, 2002, The Dynamic Volume-Return Relation of Individual Stocks, *Review of Financial Studies 15*, 1005-1047.

Lo, Andrew W., and Jiang Wang, 2000, Trading Volume: Definitions, Data Analysis, and Implications of Portfolio Theory, *Review of Financial Studies* 13, 257-300.

Lo, Andrew W., and Jiang Wang, 2006, Trading Volume: Implications of an Intertemporal Capital Asset Pricing Model, *Journal of Finance* 61, 2805-2840.

Miller, Edward, 1977, Risk, Uncertainty and Divergence of Opinion, *Journal of Finance* 32, 1151-1168.

O'Neil, William J., 1995, How To Make Money In Stocks, McGraw-Hill.

Peterson, Mitchell A., 2007, Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches, forthcoming in *Review of Financial Studies*.

Sadka, Ronnie, 2006, Momentum and Post-earnings-announcement Drift Anomalies: The Role of Liquidity Risk, *Journal of Financial Economics* 80, 309-349.

Shefrin, Hersh, and Meir Statman, 1985, The Disposition to Sell Winners too Early and Ride Losers too Long: Theory and evidence, *Journal of Finance* 40, 777-790.

Starks, Laura T., Li Yong, and Lu Zheng, 2006, Tax-Loss Selling and the January Effect: Evidence from Municipal Bond Closed-End Funds, *Journal of Finance* 61, 3049-3067.

Teo, Melvyn, and Paul O'Connell, 2007, Institutional Investors, Past Performance, and Dynamic Loss Aversion, forthcoming in *Journal of Financial and Quantitative Analysis*.

Tkac, Paula A., 1996, A Trading Volume Benchmark: Theory and Evidence, *Journal of Financial and Quantitative Analysis* 34, 89-114.

Vuolteenaho, Tuomo, 2002, What Drives Firm-level Stock Returns?, *Journal of Finance* 57, 233-264.

Wang, Jiang, 1994, A Model of Competitive Stock Trading Volume, *Journal of Political Economy* 102, 127-167.

Zhang, X. Frank, 2006, Information Uncertainty and Stock Returns, *Journal of Finance* 61, 105-137.

CHAPTER 3

AN ANALYSIS OF CAPITAL GAINS OVERHANG APPROXIMATION METHODOLOGIES

3.1 INTRODUCTION

Key to the analyses in Chapter 2 is the estimation of an aggregate Capital Gains Overhang (CGO) value for a given stock on a given day. This CGO value is calculated using a modified version of Grinblatt and Han's (2005) methodology, in which daily trading volumes are combined with daily closing prices to iteratively update a reference price, which represents the volume-weighted average purchase price of the stock. The CGO value is then calculated as the difference between this reference price and the current closing price.[1]

Due to data limitations,[2] there are two important assumptions made by this methodology. First, it assumes that sales are made by all existing shareholders in proportion to their current holdings. Thus, if 1% of outstanding shares are traded on a given day, it is assumed that all existing shareholders sell 1% of their holdings. Second, it assumes that all transactions occur at the daily closing price.

In this Chapter, I utilize a transaction-level data set from an American brokerage firm to empirically examine the validity of these two assumptions. Additionally, I examine two alternate methods of assigning sell transactions, *first-in-first-out* (FIFO) and *last-in-first-out* (LIFO), to determine if they are more accurate than the weighted average (WA) method currently used. An overview of the various CGO values calculated and the tests performed in this Chapter is shown in Figure 3.1 below.

---

[1] Full details of this methodology are given in Chapter 2.

[2] The modified version of Grinblatt and Han's (2005) methodology relies on daily market data from CRSP. (Specifically, the daily closing price and the total number of shares traded are utilized.) It is the lack of more detailed (i.e. transaction-level) data which necessitates the subsequently mentioned simplifying assumptions.

[Figure 3.1 about here]

3.2 DATA

The main data set used in this chapter consists of transaction-level data from an American brokerage firm.[3] The full data set spans the January 1, 1991 to November 29, 1996 date range, and consists of 1,854,776 transactions in 102,512 customer accounts, and includes 10,877 stocks. I supplement this data with price and share adjustment factors, closing prices, outstanding shares and exchange data from the CRSP daily data file. I then filter the combined data, removing transactions with missing or invalid transaction and/or closing price(s), those not currently trading on the NYSE or AMEX[4], and those which are traded by fewer than 10 accounts over the entire date range.[5] This results in a sample data set of 1,113,762 transactions in 90,686 customer accounts, and includes 2,831 stocks.

3.3 ACTUAL CGO VALUES BASED ON TRANSACTION PRICES

3.3.1 Introduction

I start my analysis by calculating actual CGO values for each *Date-Stock-Account* using transaction prices. (This will be referred to as Actual Transaction Price CGO, or ATP CGO). Prior to performing the calculations, two set of modifications are applied to the data. These are described below.

---

[3] The transaction-level data was generously provided to me by Professor Alok Kumar from the Department of Finance, McCombs School of Business, University of Texas at Austin.

[4] Only NYSE and AMEX stocks are selected in order to maintain comparability with the results from Chapter 2.

[5] This filter is used to ensure that thinly traded stocks, with potentially more volatile prices, do not unduly influence the results. After applying the two previous filters, this results in the removal of 932 transactions, or approximately 0.08% of the remaining data set.

3.3.2 *Share and price comparability correction*

The transaction-level data reports the actual number of shares transacted as well as the actual transaction price. In order to ensure comparability across time, these values are adjusted to account for stock splits and other such events. Similar adjustments are made to the closing price and outstanding share data sourced from the CRSP daily file. To make these modifications, the cumulative factor to adjust shares[6] and the cumulative factor to adjust price[7] from the daily CRSP file are employed. The number of shares transacted and the number of shares outstanding are both multiplied by the cumulative factor to adjust shares, and the transaction prices and daily closing prices are divided by the cumulative factor to adjust prices.

3.3.3 *Negative balance correction*

A potential problem with the transaction-level data set is that the starting date has been selected to correspond with the first trading day of a calendar year, and does not necessarily represent the opening date of investors' accounts. Thus, investors may have purchased shares prior to January 1, 1991, but sell them on or after this date. In such cases, it will appear as if investors are holding a negative inventory of shares. To correct for this, each stock in each account is examined from the start of the holding period to the end. If at any time during this period a transaction occurs that would result in a negative share balance, the number of shares sold is adjusted to result in a zero share balance.

For example, suppose that Account 1 buys 1,000 shares of Stock A on January 2, 1993, and then sells 2,000 shares of Stock A on January 3, 1993 (with no intervening transactions in Stock A). In such a situation, Account 1 is assumed to be selling 1,000 pre-existing shares (i.e. purchased before the start of the data range) and the 1,000 shares purchased on January 1, 1993 (i.e. purchased within the data range). Thus, the number of shares sold on February 1, 1993 is adjusted to 1,000, resulting in a net balance at the end

---

6 Data item CFACSHR from the CRSP daily data file.
7 Data item CFACPR from the CRSP daily data file.

of January 3, 1993 of zero shares, and the cost basis for the *sell* transaction will be based on the January 1, 1993 *buy* transaction price.

The net effect of such adjustments is to ignore any 'pre-existing' share balances for which a cost basis and CGO value cannot be accurately calculated, and only include those for which an exact CGO value can be calculated.[8] All subsequent calculations use these adjusted transaction amounts.[9]

3.3.4 Base calculation methodology

For the base calculation, my objective is to produce an aggregate cost basis (CB) and share balance (SB) at the *Date-Stock* level. These can then be used to calculate a cost per share (CPS), which can then be compared to the current market price to arrive at a CGO value. To do this, I first calculate cost basis adjustment (CBA) and share balance adjustment (SBA) values at the *Account–Stock–Date* level.[10] These are then aggregated across accounts, and the results are processed chronologically to produce the running total CB and SB values by *Date-Stock.*

SBA values for each *Date-Stock-Account* are simply the number of shares transacted (*buys* have positive values, and *sells* have negative values). CBA values for each *Date-Stock-Account* are calculated as follows.

For *buys*:

$$CBA = shares\ transacted \times transaction\ price \qquad [1]$$

8 Approximately 16.8% of transactions require this modification.

9 Another potential complication involves situations where there are multiple transactions for a given *Account–Stock-Date*. As there are no time stamps on the transactions, they are processed from largest buy (based on number of shares) to smallest buy, and then from smallest sale to largest sale. This minimizes the number of adjustments that will be required. Approximately 1.4% of transactions fall on such multiple transaction days.

10 CBA and SBA values are the amounts by which the cost basis and share balance increases/decreases for a given *Date-Stock-Account*. This differs from the CB and SB values which are the running total of all prior CBA and SBA values. In other words, the CBA and SBA represent the transaction values, while the CB and SB represent the account balances.

For *sells*:

$$CBA = \textit{shares transacted} \times \textit{historic average cost per share (CPS)} \qquad [\,2\,]$$

Thus, sales are assumed to occur in a pro-rata fashion across all existing holdings for the given *Account–Stock.* This is an admitted limitation of this methodology, as in reality an investor is able to specify the sale of specific share blocks for tax purposes. For example, an investor may find it beneficial to sell shares purchased at the highest price to minimize capital gains taxes. Alternately, if an investor has capital losses from other investments, she may want to sell shares purchased at a lower price, thus realizing a gain at a time when it can be immediately offset by a loss. Lacking information on which blocks are being sold by investors, I apply a pro-rata methodology. Thus, there are still assumptions being made regarding sales even using transaction-level data, but it has now been reduced to the account level as opposed to the overall market level as in the Grinblatt and Han (2005) methodology.

Next, CB and SB values are summed across *Accounts* by *Date-Stock* to arrive at aggregate CB and SB values at the *Date-Stock* level. A potential complication with the CB and SB data at the *Date-Stock* level is that not every stock will be transacted by at least one account on every date. To avoid gaps in the time series, I merge the above data into a second data file (extracted from the CRSP daily data file) which contains a complete date range for each stock. CB and SB values for missing days are populated using the most recent available values. Thus, for days with no trading activity, the CB and SB remain unchanged, although the CGO will fluctuate with changes in the daily closing price (as described below).

The final step is to calculate actual CGO values by *Date-Stock.* To do this, the cost per share (CPS) is first calculated as follows:

$$CPS_{s,t} = \frac{CB_{s,t}}{SB_{s,t}} \qquad [\,3\,]$$

where $CPS_{s,t}$ is cost per share for stock s on date t, $CB_{s,t}$ is the cost basis for stock s on date t, and $SB_{s,t}$ is the share balance for stock s on date t. CGO is then calculated as follows:

$$CGO_{s,t} = \frac{(P_{s,t} - CPS_{s,t})}{P_{s,t}} \qquad [\ 4\ ]$$

where $CCO_{s,t}$ is the capital gains overhang for stock s on date t, $P_{s,t}$ is the daily CRSP closing price for stock s on date t, and $CPS_{s,t}$ is cost per share for stock s on date t as defined in [3].

This is similar to the methodology used by Grinblatt and Han (2005), in that the current price is used as the denominator in the equation. Mathematically, this will increase the magnitude of negative CGO values (i.e. capital losses) in relation to positive CGO values (i.e. capital gains), as noted in Chapter 2. A difference compared to the methodology of Chapter 2 is that I am not using lagged values (Chapter 2 uses values from date *t-5*). This lagging is not necessary, as I am not performing an event time analysis, but I am comparing CGO values on a specific date calculated using different methodologies. Thus, an 'event' that impacts a stock on a given date will influence all of the methodologies concurrently.

A detailed example of this calculation methodology can be found in Appendix 3.1, Panel A.

3.3.5 Base case results

Figure 3.2 shows the mean ATP CGO values by date, and Figure 3.3 shows the 10^{th}, 50^{th}, and 90^{th} percentile values of the ATP CGO distribution by date.[11]

[Figure 3.2 about here]

[11] All CGO values, including those in subsequent sections, are trimmed at 1% / 99% by date. This will reduce the impact of large outliers, which is of particular concern for negative CGO values which can range to $-\infty$.

For Figure 3.2, an equal-weighted average CGO value (across stocks) is calculated for each *Date*. On average, the mean value for each date is based on CGO values for 2,200 stocks (ranging from a low of 152 stocks to a high of 2,380 stocks), and 2.05 accounts for each stock (ranging from a low of 1 account to a high of 276 accounts).

[Figure 3.3 about here]

There are two key observations to be made regarding Figure 3.3. First, the magnitude of negative CGO values is typically greater than the magnitude of positive CGO values. This is a result of using the current price in the denominator of the CGO calculation, which increases the magnitude of negative values in relation to positive values.[12] Second, the dispersion of CGO values increases over time. This is to be expected, as the actual CGO calculation methodology does not incorporate a calibration period as in Chapter 2.[13] The lack of a calibration period is not a major concern, though, as I am not attempting to construct an exact measure of CGO for the entire market, but rather I am attempting to compare CGO values constructed using different methodologies. Even so, to avoid potential problems resulting from thinly traded stocks at the start of the data set, I introduce two measures. First, as previously mentioned, only stocks which are traded within 10 or more accounts across the entire date range are included in the sample. Second, in my analysis I will focus on paired-sample *t*-tests which will compare results on a daily basis, which will minimize issues related to the drift of CGO values over time.

---

12 Based on Formula [4], possible CGO values range from +1 to $-\infty$.

13 In the modified version of Grinblatt and Han's (2005) methodology used in Chapter 2, a calibration period of five years was used before including a stock in the sample set. This was done to mitigate the impact of the initial assumption that 100% of outstanding shares traded at the first day's closing price. No such assumption is made in the actual method used in this Chapter, so a calibration period is not used.

3.4 ACTUAL CGO BASED ON DAILY CLOSING PRICES

3.4.1 Introduction

One of the assumptions in the Grinblatt and Han (2005) methodology is that all transactions, whether *buys* or *sells*, occur at the day's closing price. Thus, my first test involves recalculating the CGO values from the prior section using the daily closing price for all *buy* and *sell* transactions instead of the actual transaction price. (This will be referred to as the Actual Closing Price CGO, or ACP CGO.) Comparing the results to the base case above will provide insight into the reasonableness of this assumption.

3.4.2 ACP CGO results

The mean of daily CGO values calculated using transaction prices is -0.005957 while that based on closing prices is -0.004816. Figure 3.4 plots the difference between the two CGO values on daily basis.

[Figure 3.4 about here]

3.4.3 Paired t-tests by date

With a few exceptions early in the date range, the ACP CGO values are higher than ATP CGO values. To further test the significance of these differences, I run a paired-sample *t*-test which compares the daily mean values for both methods (i.e. 1,518 pairs of daily values are used). The resultant p-value is 0.0000, indicating that there is a significant difference in the daily mean values at all conventional significance levels.

3.4.4 Paired t-tests by date–stock

As a further check, I run a similar paired *t*-test, but at the stock level. Thus, for each stock, I compare 1,518 pairs of daily CGO values. The average *p*-value for this stock-level paired *t*-test is 0.01855. The breakdown of p-values is shown in Table 3.1 below.

[Table 3.1 about here]

As can be seen, the vast majority of stocks have highly significant differences in CGO values.

3.4.5 CGO tercile / quintile migration

So far, there is strong evidence that using the daily closing price instead of the actual transaction price does result in statistically different CGO values. While this is important, even more important is whether these differences occur uniformly across the various stocks. This is critical as the analyses in Chapter 2 divide the data into CGO terciles/quintiles and examine the relative performance of the various groups. Thus, if using closing prices merely increases all CGO values by a fixed amount, then the stocks' relative CGO values and resultant terciles/quintiles will not be changed.

To test this, I assign stocks into CGO terciles (quintiles) on a daily basis using ATP CGO values and breakpoints calculated using contemporaneous daily ATP CGO values. I then assign the same stocks to CGO terciles (quintiles) on a daily basis using ACP CGO values and breakpoints calculated using contemporaneous daily ACP CGO values. I then construct a transition matrix, to see how many *Date–Stock* observations fall into a different CGO tercile (quintile) using the different methods. Results are shown in Table 3.2 below. Panel A shows tercile-based results, and Panel B shows quintile-based results.

[Table 3.2 about here]

By summing the diagonals (in bold), I can determine how many observations stay in the same tercile (quintile) across the two assignment methods. I find that 97.5% (95.5%) of observations stay in the same tercile (quintile). Only 0.0% (0.1%) of observations shift more than 1 tercile (quintile). Additionally, the percent of observations which do shift appears to be relatively evenly distributed across terciles (quintiles), and does not appear to be more prevalent in any specific tercile (quintile).

Thus, it appears as if the impact of using closing prices, while significant, is applied relatively uniformly across the terciles/quintiles and does not cause an inordinate number of observations to be reclassified into a different CGO tercile or quintile.

3.5 WEIGHTED AVERAGE, FIFO AND LIFO METHODS

3.5.1 Introduction

The prior section concluded that using closing prices instead of transaction prices does introduce a bias to the CGO values, but that this bias does not result in an unacceptable number of observations being misclassified regarding CGO terciles or quintiles. Building on this result, the next step is to evaluate three potential methodologies for calculating CGO values using closing prices: weighted average (WA), first-in-first-out (FIFO), and last-in-first-out (LIFO). (These will be referred to as WA CGO, FIFO CGO and LIFO CGO, respectively.)

3.5.2 Methodology

For all three of these methods, *buys* and *sells* are aggregated at the *Date–Stock* level (i.e. *buys* and *sells* are summed across all accounts), and the summed transaction is assumed to occur at the daily closing price.

For *buys*, all three methods use the same calculation. Specifically, the cost basis for a stock is increased by the number of shares purchased multiplied by the daily closing price. For *sells*, each of the methods uses a different calculation. For the weighted average method, it is assumed that all existing shareholders sell an equal percentage of their holdings. Thus, the cost per share remains unchanged, and only the total number of shares held changes (i.e. decreases). For the FIFO method, it is assumed that the first purchasers sell first (i.e. shares purchased on the earliest date are sold first). Thus, the total cost basis is reduced by the number of shares sold multiplied by the purchase price of the earliest purchaser.[14] For the LIFO method, it is assumed that the last purchaser sells first (i.e. shares purchased on the latest date are sold first). Thus, the total cost basis is reduced by the number of shares sold multiplied by the purchase price of the last purchaser.[15]

Detailed examples of the various calculation methods are shown in Appendix 3.2. Panel B provides an example of the WA CGO calculation (which is identical to the ACP CGO calculation in this case). Panel C provides an example of the FIFO CGO calculation, and Panel D provides an example of the LIFO CGO calculation.

3.5.3 WA CGO, FIFO CGO and LIFO CGO results

Results for these three new methods, as well as the base case (ACP CGO) method, are shown in Table 3.3 below.

[Table 3.3 about here]

Panel A of Table 3.3 shows the means and standard deviations of the daily averages (equally-weighted across all stocks) of the various CGO values. Panel B of Table 3.3

[14] If more shares are sold than were purchased by the first purchaser, the excess is assumed to be sold by the second remaining purchaser. Any excess shares from this transaction are assumed to be sold by the third remaining purchaser, and so on.

[15] If more shares are sold than were purchased by the last purchaser, the excess is assumed to be sold by the second last remaining purchaser. Any excess shares from this transaction are assumed to be sold by the third last remaining purchaser, and so on.

shows the means and standard deviations of the daily differences between the three new CGO values and the ACP CGO value, respectively. As can be seen, the WA CGO method produces slightly lower values, while the FIFO and LIFO methods produce higher values. This is particularly true for the LIFO method, although this is not surprising as the LIFO method produces the largest CGO values in appreciating markets.[16] Based on the magnitude of the difference in means, the FIFO method appears best, but based on the standard deviation of the difference in means the WA method appears best. Thus, the results so far are not conclusive.

To get a better understanding of these results, I plot the daily differences for the above three pairs in Figure 3.5 below. Panel A shows the difference between the WA CGO and the ACP CGO, Panel B shows the difference between the FIFO CGO and the ACP CGO, and Panel C shows the difference between the LIFO CGO and the ACP CGO.

[Figure 3.5 about here]

Consistent with the standard deviations reported in Panel B of Table 3.3, it can be seen that the differences are least volatile for the WA CGO values, and increasingly volatile for the LIFO and FIFO CGO values, respectively. Also, it can be seen in Panel C that LIFO CGO values are almost always greater than those of the base case (ACP CGO), as previously noted.

3.5.4 Paired t-tests by date

Following the analysis methods of Section 3.4, I next calculate *p*-values for paired *t*-tests by date for each of the combinations (WA CGO vs. ACP CGO, FIFO CGO vs. ACP CGO and LIFO CGO vs. ACP CGO). In all cases, the p-values are 0.000, indicating that there is a significant difference in the daily mean values at all conventional significance levels. Thus, this test sheds little additional light on which method is best.

[16] This is illustrated in Panel D of Appendix 3.1.

3.5.5 *Paired t-tests by date–stock*

Next, I run similar paired *t*-tests, but at the stock level. Thus, for each stock, I compare 1,518 pairs of daily CGO values. This is performed once for each of the three comparison sets described above. The breakdown of *p*-values is shown in Table 3.4 below.

[Table 3.4 about here]

The results are remarkably consistent across the CGO calculation methods, with only a handful of stocks not having a significant difference. Thus, the tests so far merely validate that all three methods produce CGO values that are significantly different than the ACP CGO values, but they do not definitively indicate any one method is superior.

3.5.6 *CGO tercile / quintile migration*

As a final test, I construct tercile and quintile migration matrices for the various methods in a manner similar to that in Section 3.4.5. Results are shown in Table 3.5 below. Panel A shows tercile-based results, and Panel B shows quintile-based results.

[Table 3.5 about here]

Examining Panel A of Table 3.5, it can be seen that the WA CGO method has the largest percentage of observations in the same CGO tercile (89.1%), as well as the smallest percentage of observations shifting more than one CGO tercile (0.3%). Interestingly, the LIFO method is second-best, clearly beating the FIFO method on these metrics.

Panel B of Table 3.5 shows that results are similar for CGO quintiles. Specifically, the WA CGO method is clearly the best, with the LIFO CGO method being second-best and the FIFO CGO method being third-best.

Thus, the WA CGO method produces the lowest amount of misclassification errors for both CGO terciles and quintiles. Even so, it should be noted that 10.9% (18.9%) of CGO terciles (quintiles) are being misclassified using this method. While the weighted average method is the best of the three available methods, it still leaves room for improvement.

3.6 CONCLUSION

In this Chapter, I analyze the accuracy of the modified Grinblatt and Han (2005) capital gains overhang (CGO) computation methodology employed in Chapter 2. This is done by utilizing a transaction-level data set from an American brokerage firm, which allows for the computation of CGO values accurate to the level of individual investor accounts.

The first major assumption employed in the modified Grinblatt and Han (2005) methodology is the use of daily closing prices instead of actual transaction prices in the CGO calculations. Thus, my first test involves comparing CGO values based on actual transaction prices (ATP CGO) with those based on daily closing prices (ACP CGO) in order to check the reasonableness of this assumption. I find that a bias is introduced by this assumption, increasing the daily CGO values by a highly significant 0.11%. I further examine the impact of this bias on the allocation of observations to CGO terciles (quintiles), and find that the impact is uniformly spread across observations and that only 2.5% (4.5%) of observations are categorized into incorrect terciles (quintiles) as a result.

The second major assumption tested is the use of a weighted average method for calculating the impact of *sell* transactions on the total cost basis and resultant CGO values (WA CGO). Under this method, all existing shareholders are assumed to sell an equal percentage of their existing holdings. This method is compared to a *first-in-first-out* (FIFO) method in which the earliest purchasers are assumed to sell first, and a *last-in-first-out* (LIFO) method in which the latest purchasers are assumed to sell first. All three methods result in CGO values that have a statistically significant difference from the ACP CGO values, but based on the number of observations being misclassified into CGO terciles (quintiles) the weighted average method is clearly superior to the FIFO and LIFO methods. Even so, approximately 10.9% (18.9%) of observations are classified into incorrect CGO terciles (quintiles) with the weighted average method.

An important future addition to this work would be to determine if there are potential adjustments to the WA CGO calculation methodology which would minimize the number of observations that are misclassified by CGO terciles/quintiles. One possible way of achieving this would be to determine if the over/underestimation of CGO values using the weighted average methodology are systematically related to specific stock characteristics (for example size, B/M, or even industry sectors), and then apply relevant 'adjustment' factors to CGO values for such stocks.

Table 3.1 – P-value distribution for equality of means

Significance Level	Number of Stocks	Percent of Stocks
Not significant	109	4%
10% Level	22	1%
5% Level	25	1%
2% Level	2,674	94%
Totals	**2,830**	**100%**

Table 3.2 – CGO tercile and quintile transition matrices (ATP CGO versus ACP CGO)

Panel A – Tercile Transition Matrix

	ACP CGO Tercile		
ATP CGO Tercile	**1**	**2**	**3**
1	**32.4%**	0.7%	0.0%
2	0.7%	**32.8%**	0.6%
3	0.0%	0.6%	**32.4%**

Same Tercile:	**97.5%**
+/- 1 Tercile:	2.5%
Other:	0.0%

Panel B – Quintile Transition Matrix

	ACP CGO Quintile				
ATP CGO Quintile	**1**	**2**	**3**	**4**	**5**
1	**19.5%**	0.5%	0.0%	0.0%	0.0%
2	0.5%	**18.8%**	0.7%	0.0%	0.0%
3	0.0%	0.7%	**18.7%**	0.6%	0.0%
4	0.0%	0.0%	0.6%	**18.9%**	0.4%
5	0.0%	0.0%	0.0%	0.4%	**19.5%**

Same Quintile:	**95.5%**
+/- 1 Quintile:	4.4%
Other:	0.1%

Table 3.3 – Summary statistics for CGO calculation methodologies

Panel A – Means and Standard Deviations for CGO Values

	Mean	Std Dev
ACP CGO	-0.004816	0.048063
WA CGO	-0.006790	0.047674
CGO FIFO	-0.003724	0.048156
CGO LIFO	0.003998	0.049375

Panel B – Means and Standard Deviations for Differences in CGO Values

	Mean	Std Dev
WA CGO - ACP CGO	-0.001974	0.003554
FIFO CGO - ACP CGO	0.001092	0.006318
LIFO CGO - ACP CGO	0.008814	0.004918

Table 3.4 – P-value distributions for equality of means

	Weighted-Average		FIFO		LIFO	
Significance Level	Number of Stocks	Percent of Stocks	Number of Stocks	Percent of Stocks	Number of Stocks	Percent of Stocks
Not significant	77	2.72%	73	2.58%	66	2.33%
10% Level	14	0.49%	14	0.49%	17	0.60%
5% Level	18	0.64%	16	0.57%	15	0.53%
2% Level	2,721	96.15%	2,727	96.36%	2,732	96.54%
Totals	**2,830**	**100.00%**	**2,830**	**100.00%**	**2,830**	**100.00%**

Table 3.5 – CGO tercile and quintile transition matrices

Panel A – Tercile Transition Matrix

	WA CGO Tercile			**FIFO CGO Tercile**			**LIFO CGO Tercile**		
ACP CGO Tercile	**1**	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**
1	**30.6%**	2.3%	0.1%	**28.1%**	4.1%	0.8%	**29.7%**	2.9%	0.4%
2	2.3%	**28.8%**	3.0%	4.2%	**24.9%**	5.1%	3.2%	**26.9%**	4.1%
3	0.1%	3.0%	**29.7%**	0.7%	5.1%	**26.9%**	0.1%	4.4%	**28.2%**

Same Tercile:	**89.1%**	Same Tercile:	**79.9%**	Same Tercile:	**84.8%**
+/- 1 Tercile:	10.6%	+/- 1 Tercile:	18.6%	+/- 1 Tercile:	14.6%
Other:	0.3%	Other:	1.5%	Other:	0.6%

Panel B – Quintile Transition Matrix

	WA CGO Quintile					FIFO CGO Quintile					LIFO CGO Quintile				
ACP CGO Quintile	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**
1	**18.2%**	1.6%	0.1%	0.1%	0.0%	**16.3%**	2.5%	0.6%	0.3%	0.2%	**17.6%**	1.9%	0.2%	0.1%	0.1%
2	1.6%	**16.0%**	2.2%	0.2%	0.1%	2.8%	**13.0%**	3.2%	0.8%	0.3%	2.0%	**14.6%**	2.6%	0.6%	0.3%
3	0.1%	2.2%	**15.0%**	2.6%	0.1%	0.5%	3.2%	**12.0%**	3.8%	0.6%	0.2%	3.2%	**13.2%**	2.8%	0.7%
4	0.0%	0.3%	2.3%	**14.8%**	2.5%	0.2%	1.0%	3.3%	**11.8%**	3.8%	0.1%	0.3%	3.9%	**12.7%**	3.1%
5	0.0%	0.1%	0.3%	2.3%	**17.0%**	0.1%	0.4%	1.0%	3.3%	**14.9%**	0.0%	0.1%	0.2%	3.8%	**15.7%**

WA	
Same Quintile:	**81.0%**
+/- 1 Quintile:	17.4%
Other:	1.5%

FIFO	
Same Quintile:	**68.0%**
+/- 1 Quintile:	26.0%
Other:	6.1%

LIFO	
Same Quintile:	**73.9%**
+/- 1 Quintile:	23.3%
Other:	2.8%

Figure 3.1 – Overview of CGO calculations and tests

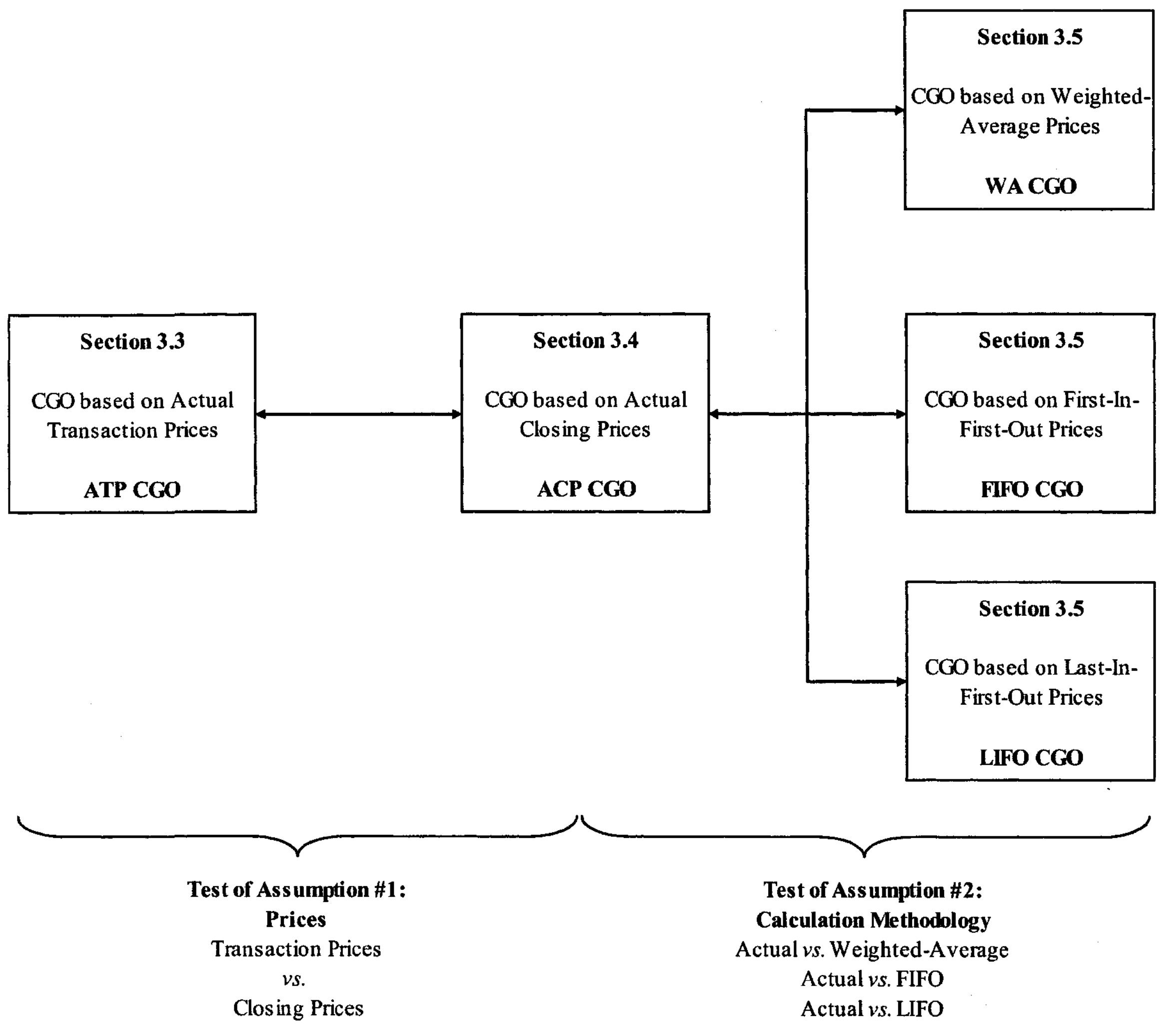


Figure 3.2 – Mean actual transaction price capital gains overhang (ATP CGO) values by date

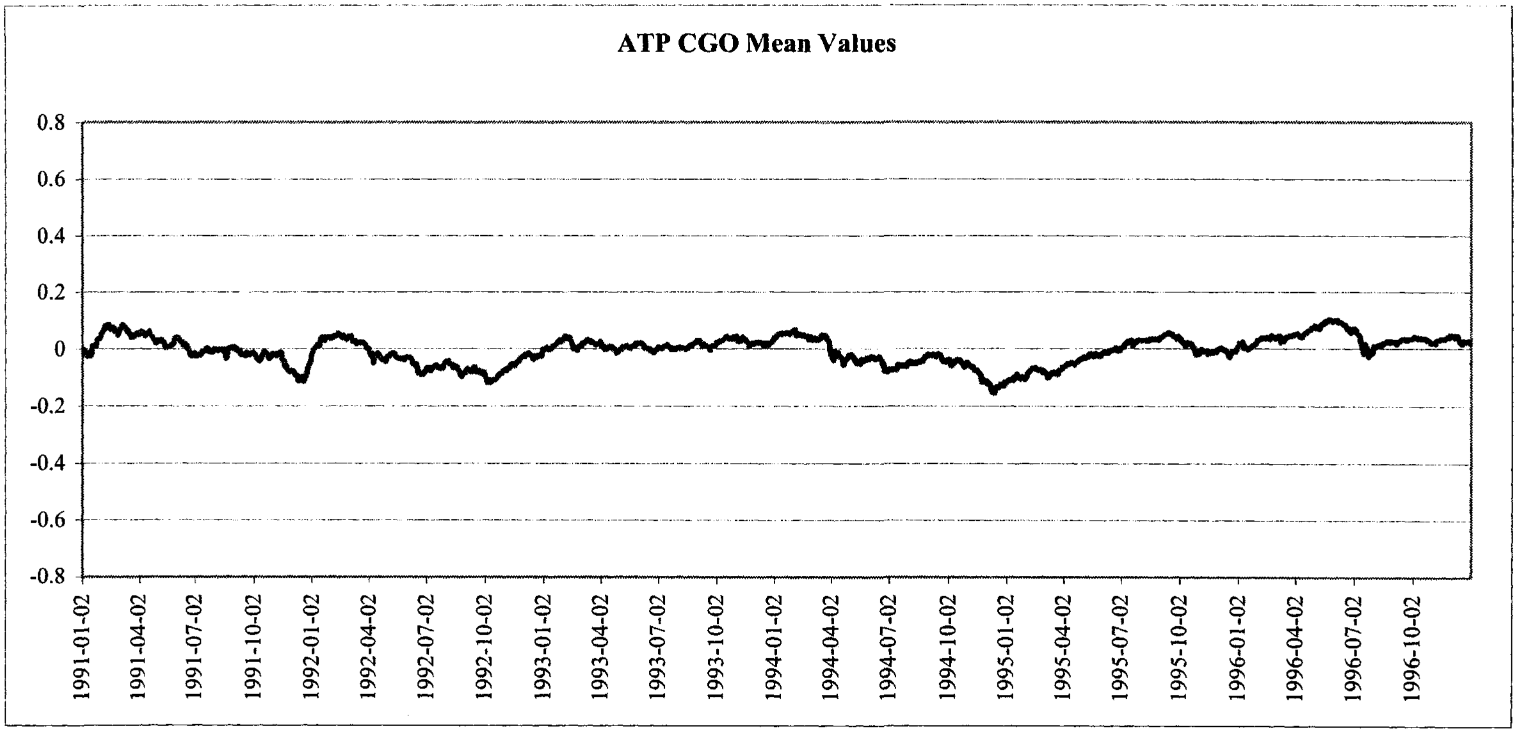


Figure 3.3 – Actual transaction price capital gains overhang (ATP CGO) percentiles by date

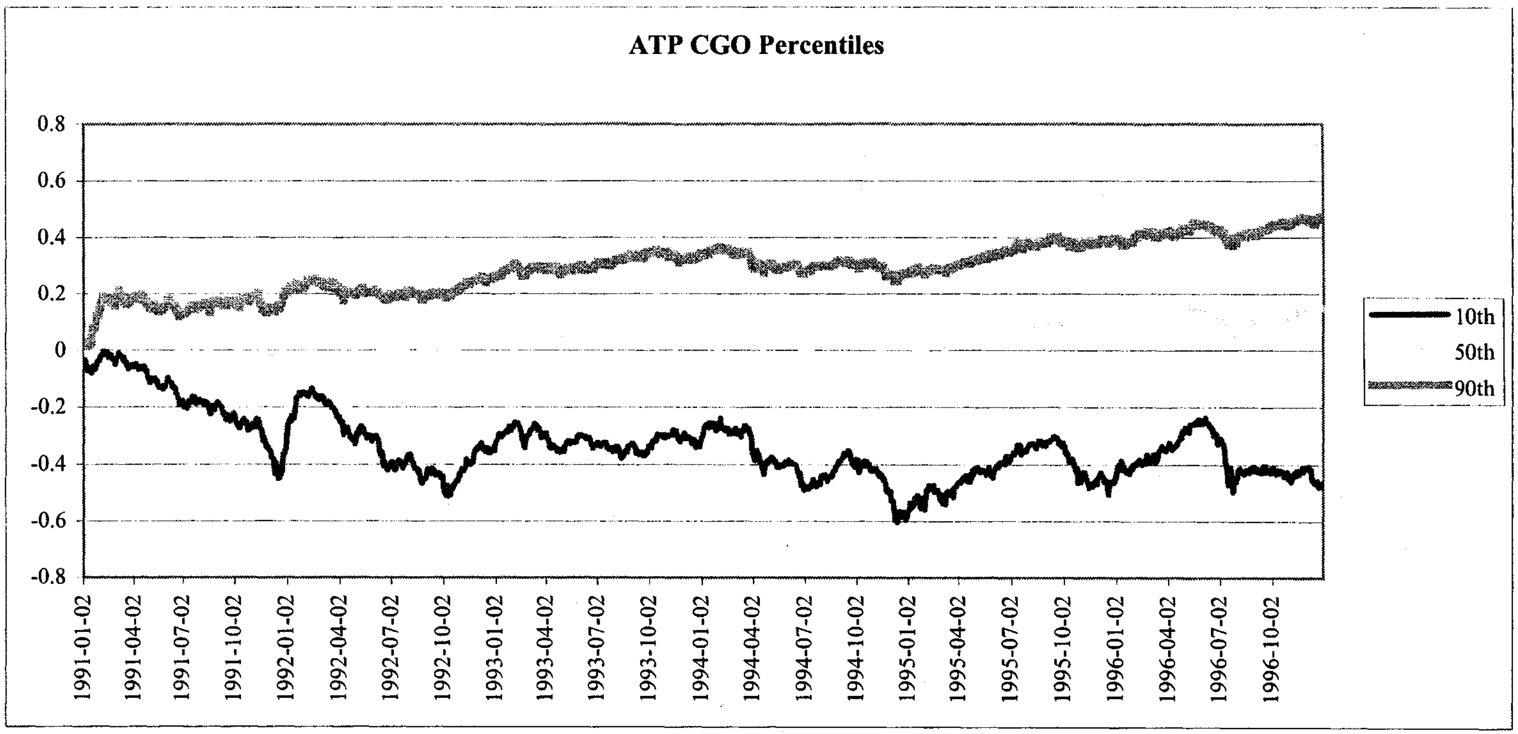


Figure 3.4 – Differences in mean CGO values by date (ACP CGO minus ATP CGO)

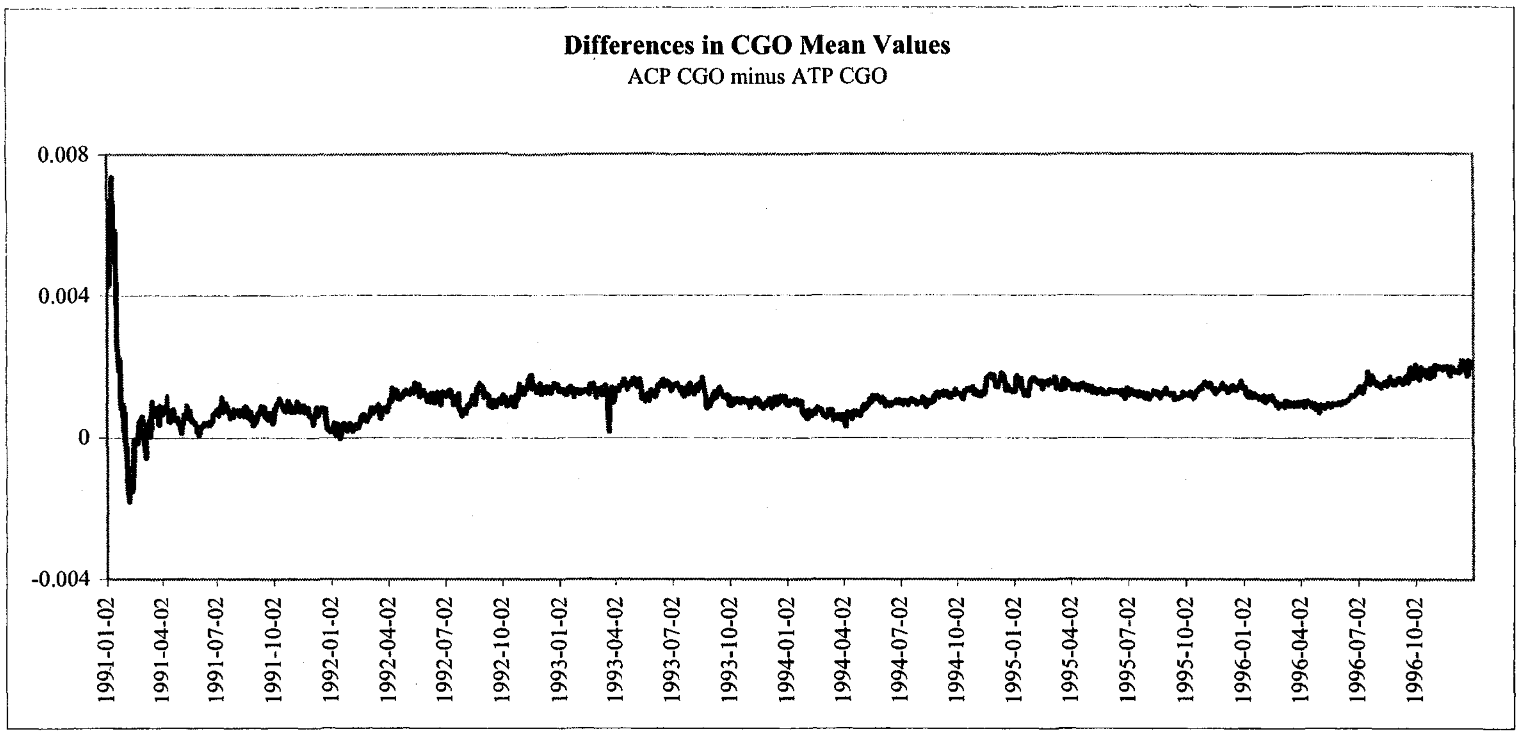
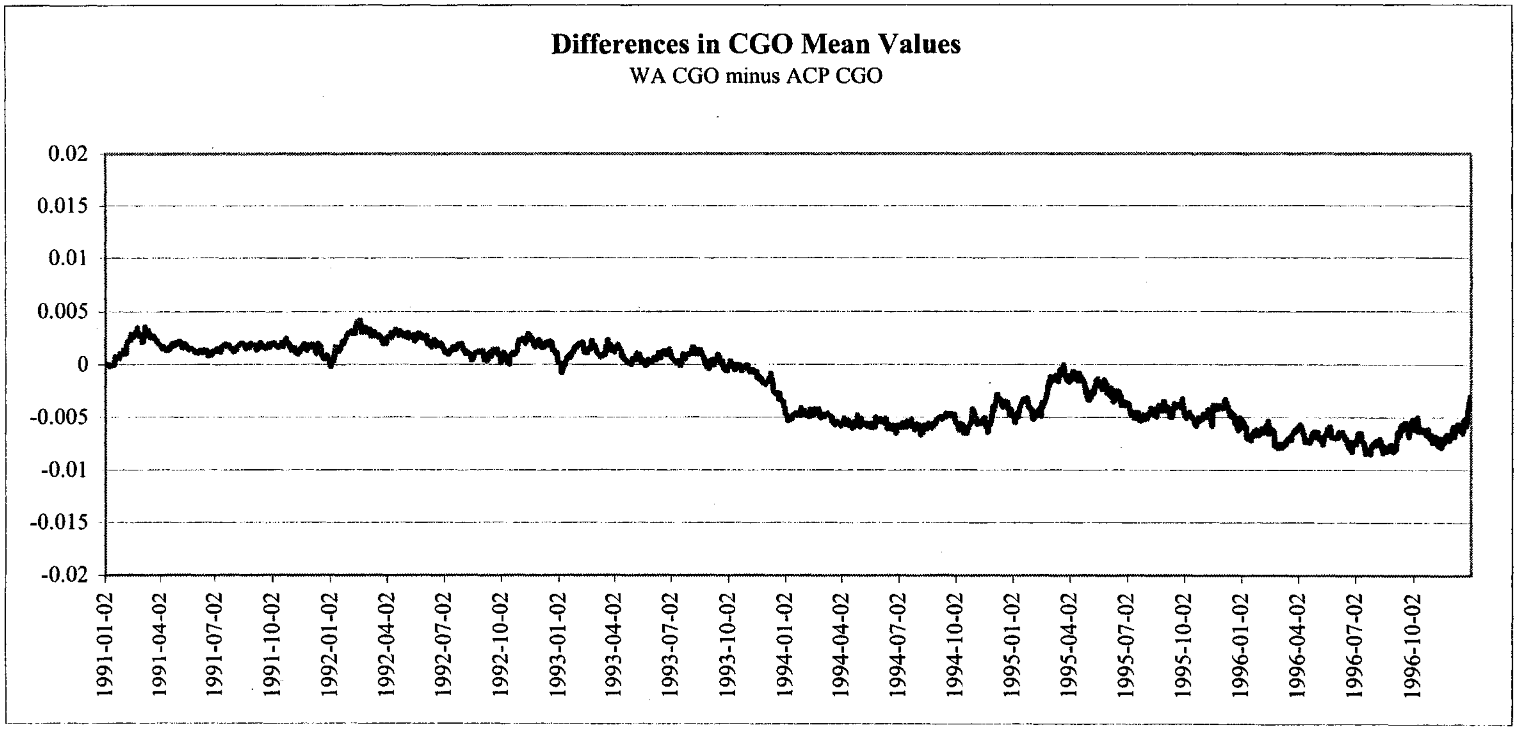
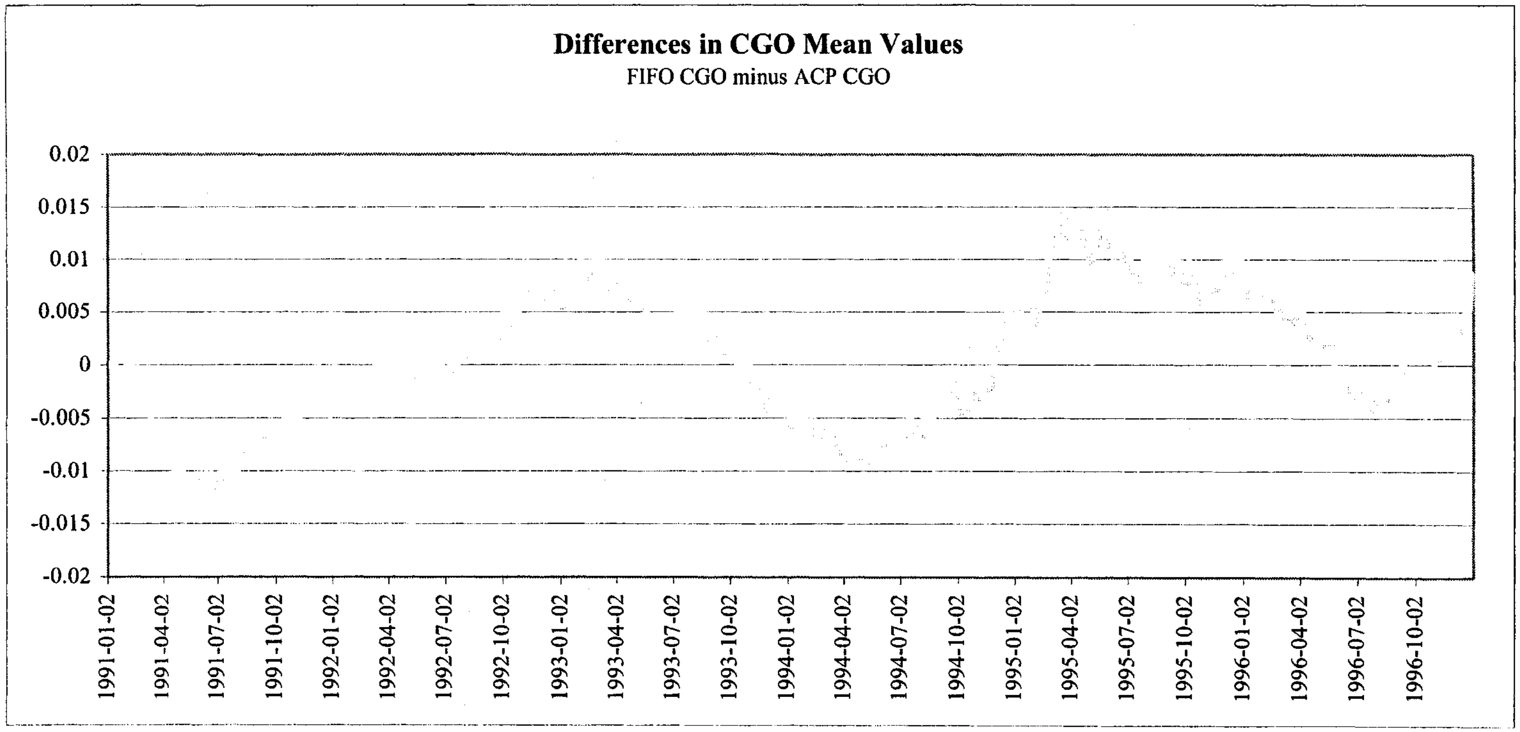


Figure 3.5 – Differences in mean CGO values by date (WA / FIFO / LIFO CGO minus ACP CGO)

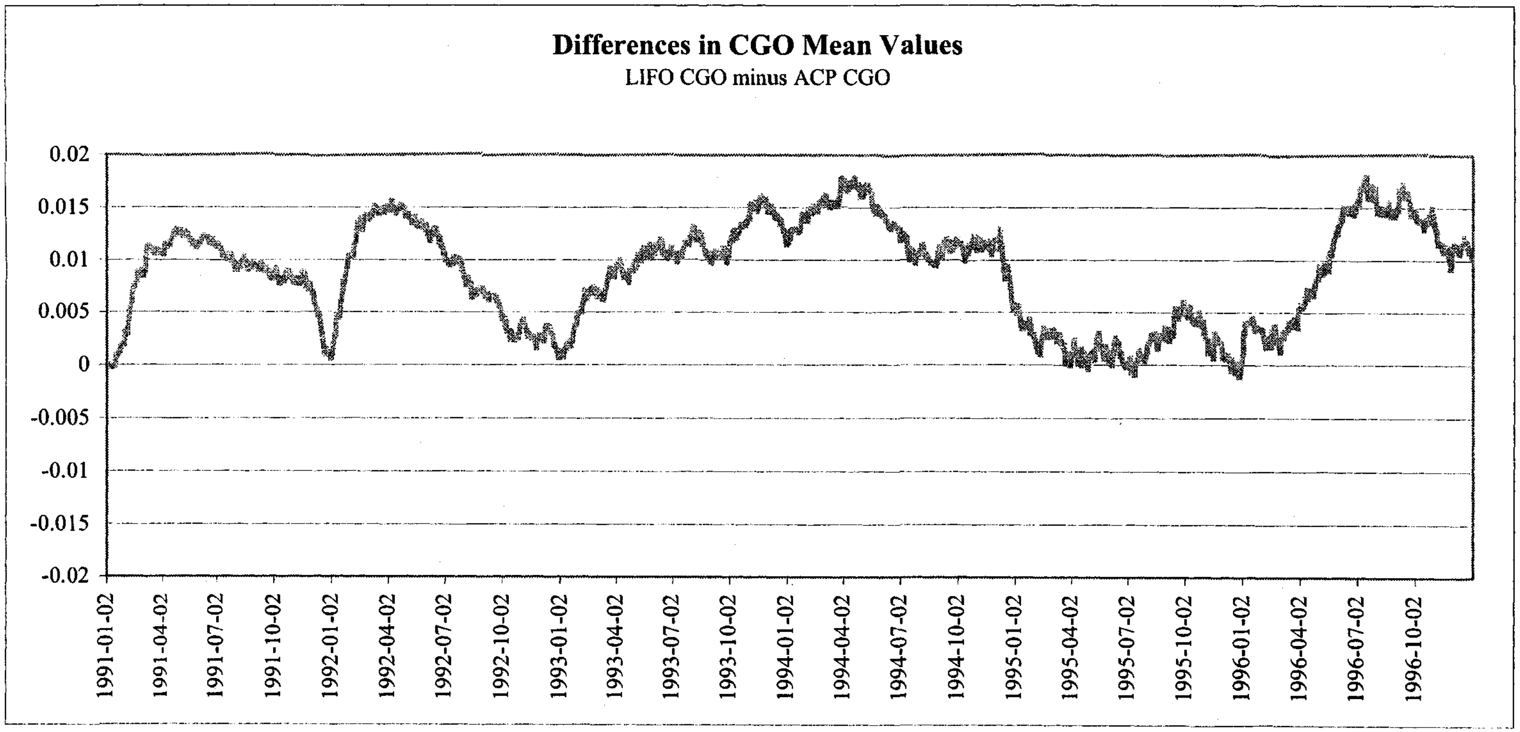
Panel A – WA CGO minus ACP CGO



Panel B –FIFO CGO minus ACP CGO



Panel C –FIFO CGO minus ACP CGO



APPENDIX 3.1 – Examples of CGO Calculation Methodologies

This Appendix gives examples of the various CGO calculation methodologies at the *Date–Stock–Account* level. Panel A shows results for Actual CGO values based on transaction prices (ATP CGO). Panel B shows results for Actual CGO values based on closing prices (ACP CGO), which, at the *Date–Account–Stock* level, will be identical to the Weighted Average CGO values based on closing prices (WA CGO). Panel C shows results for first-in-first-out CGO values (FIFO CGO). Panel D shows results for last-in-first-out CGO values (LIFO CGO).

Panel A – Example of Actual CGO Values Based on Transaction Prices (ATP CGO)

Date	Txn Shares	Txn Price	Closing Price	SBA	CBA Price	CBA	SB	CB	CPS	CGO
Jan 1, 1991	-	-	-	-	-	-	-	-	-	-
Jan 2, 1991	1,000	$ 10.00	$ 9.00	1,000	$ 10.00	$ 10,000	1,000	$ 10,000	$ 10.00	(11.11%)
Jan 3, 1991	2,000	$ 20.00	$ 19.00	2,000	$ 20.00	$ 40,000	3,000	$ 50,000	$ 16.67	12.28%
Jan 4, 1991	(500)	$ 25.00	$ 24.00	(500)	$ 16.67	$ (8,333)	2,500	$ 41,667	$ 16.67	30.56%
Jan 7, 1991	(1,000)	$ 30.00	$ 29.00	(1,000)	$ 16.67	$ (16,667)	1,500	$ 25,000	$ 16.67	42.53%
Jan 8, 1991	(1,000)	$ 15.00	$ 14.00	(1,000)	$ 16.67	$ (16,667)	500	$ 8,333	$ 16.67	(19.05%)

Variables are defined as follows:

- **Date** is the date of the transaction.
- **Txn Shares** is the number of shares bought (sold) in the transaction.
- **Txn Price** is the actual transaction price at which the shares are bought or sold.
- **Closing Price** is the daily closing price. (In this example, it is simply set to $1 less than the Txn Price, although in practice it will vary both above and below the Txn Price.)

- **SBA** is the Share Balance Adjustment, or the number of shares by which the total holdings increases (decreases) as a result of this transaction.
- **CBA Price** is the Cost Basis Adjustment Price, or the price to be used in calculating the Cost Basis Adjustment as a result of this transaction. For *buy* transactions this will always equal the Txn Price. For *sell* transactions this will always equal the most recent Cost Per Share (CPS) value.
- **CBA** is the Cost Basis Adjustment, or the total dollar amount by which the Cost Basis must be adjusted as a result of this transaction. *Buys* will be positive amounts representing an increase to the Cost Basis, and *sells* will be negative amounts representing a decrease to the Cost Basis.
- **SB** is the Share Balance, or the running total of Share Balance Adjustments for all transactions up to and including the current transaction.
- **CB** is the Cost Basis, or the running total of Cost Basis Adjustments for all transactions up to and including the current transaction.
- **CPS** is the average cost per share, which is calculated as CB / SB.
- **CGO** is the current Capital Gains Overhang, which is calculated as (Closing Price – CPS) / Closing Price.

Panel B – Example of Actual CGO Values Based on Closing Prices (ACP CGO) or Weighted Average CGO Values (WA CGO)

Date	Txn Shares	Txn Price	Closing Price	SBA	CBA Price	CBA	SB	CB	CPS	CGO
Jan 1, 1991	-	-	-	-	-	-	-	-	-	-
Jan 2, 1991	1,000	$ 10.00	$ 9.00	1,000	$ 9.00	$ 9,000	1,000	$ 9,000	$ 9.00	0.00%
Jan 3, 1991	2,000	$ 20.00	$ 19.00	2,000	$ 19.00	$ 38,000	3,000	$ 47,000	$ 15.67	17.54%
Jan 4, 1991	(500)	$ 25.00	$ 24.00	(500)	$ 15.67	$ (7,833)	2,500	$ 39,167	$ 15.67	34.72%
Jan 7, 1991	(1,000)	$ 30.00	$ 29.00	(1,000)	$ 15.67	$ (15,667)	1,500	$ 23,500	$ 15.67	45.98%
Jan 8, 1991	(1,000)	$ 15.00	$ 14.00	(1,000)	$ 15.67	$ (15,667)	500	$ 7,833	$ 15.67	(11.90%)

Variables are as defined in Panel A, with the following exception:

- **CBA Price** is the Cost Basis Adjustment Price, or the price used in calculating the Cost Basis Adjustment (CBA) as a result of this transaction. For *buy* transactions this will always equal the Closing Price. For *sell* transactions this will always equal the most recent prior Cost Per Share (CPS) value.

Panel C – Example of First-In-First-Out CGO Values (FIFO CGO)

Date	Txn Shares	Txn Price	Closing Price	SBA	CBA Price	CBA	SB	CB	CPS	CGO
Jan 1, 1991	-	-	-	-	-	-	-	-	-	-
Jan 2, 1991	1,000	$ 10.00	$ 9.00	1,000	$ 9.00	$ 9,000	1,000	$ 9,000	$ 9.00	10.00%
Jan 3, 1991	2,000	$ 20.00	$ 19.00	2,000	$ 19.00	$ 38,000	3,000	$ 47,000	$ 15.67	17.54%
Jan 4, 1991	(500)	$ 25.00	$ 24.00	(500)	$ 9.00	($4,500)	2,500	$ 42,500	$ 17.00	29.17%
Jan 7, 1991	(1,000)	$ 30.00	$ 29.00	(1,000)	$ 14.00	($14,000)	1,500	$ 28,500	$ 19.00	34.48%
Jan 8, 1991	(1,000)	$ 15.00	$ 14.00	(1,000)	$ 19.00	($19,000)	500	$ 9,500	$ 19.00	(35.71%)

Variables are as defined in Panel A, with the following exception:

- **CBA Price** is the Cost Basis Adjustment Price, or the price to be used in calculating the Cost Basis Adjustment (CBA) as a result of this transaction. For *buy* transactions this will always equal the Closing Price. For *sell* transactions this will equal the Closing Price of the earliest transaction for which the shares have not already been sold. If the earliest transaction does not have enough shares to accommodate the entire *sell* transaction, the surplus will be sold from the second earliest unused transaction, and so on until all the entire *sell* transaction has been accommodated. In such cases the CBA Price will be the share-weighted average of the respective Closing Prices.

Panel D – Example of Last-In-First-Out CGO Values (LIFO CGO)

Date	Txn Shares	Txn Price	Closing Price	SBA	CBA Price	CBA	SB	CB	CPS	CGO
Jan 1, 1991	-	-	-	-	-	-	-	-	-	-
Jan 2, 1991	1,000	$ 10.00	$ 9.00	1,000	$ 9.00	$ 9,000	1,000	$ 9,000	$ 9.00	10.00%
Jan 3, 1991	2,000	$ 20.00	$ 19.00	2,000	$ 19.00	$ 38,000	3,000	$ 47,000	$ 15.67	17.54%
Jan 4, 1991	(500)	$ 25.00	$ 24.00	(500)	$ 19.00	($9,500)	2,500	$ 37,500	$ 15.00	37.50%
Jan 7, 1991	(1,000)	$ 30.00	$ 29.00	(1,000)	$ 19.00	($19,000)	1,500	$ 18,500	$ 12.33	57.47%
Jan 8, 1991	(1,000)	$ 15.00	$ 14.00	(1,000)	$ 14.00	($14,000)	500	$ 4,500	$ 9.00	35.71%

Variables are defined in Panel A, with the following exceptions:

- **CBA Price** is the Cost Basis Adjustment Price, or the price to be used in calculating the Cost Basis Adjustment (CBA) as a result of this transaction. For *buy* transactions this will always equal the Closing Price. For *sell* transactions this will equal the Closing Price of the latest transaction for which the shares have not already been sold. If the latest transaction does not have enough shares to accommodate the entire *sell* transaction, the surplus will be sold from the second latest unused transaction, and so on until all the entire *sell* transaction has been accommodated. In such cases the CBA Price will be the share-weighted average of the respective Closing Prices.

BIBLIOGRAPHY

Grinblatt, Mark, and Bing Han, 2005, Prospect Theory, Mental Accounting, and Momentum, *Journal of Financial Economics* 78, 311-339.

CHAPTER 4

THE IMPACT OF DROPPED COVERAGE ON ANALYSTS' CONSENSUS RECOMMENDATIONS AND EARNINGS PER SHARE ESTIMATES

4.1 INTRODUCTION

Virtually all brokerage and financial management firms employ stock market analysts. Their functions range from global macroeconomic analysis to company level analysis. In this paper, my interest lies with the company level analysts. Typically, such analysts will specialize in one or more industry sectors, such as automotives or retailing. Within a sector, a given analyst will typically follow a set of companies, for which he will produce periodic comprehensive research reports supplemented by more frequent, but less detailed, updates. Within the research report, the analyst will usually give a qualitative overview of the company's current operations and future prospects, provide pro-forma financial statements or estimates of key financial numbers (including earnings) for the next one to three years (the *estimate*), and give their final opinion on whether the stock should be purchased or sold (the *recommendation*).

Analysts' reports and updates are heavily relied upon by both institutional and individual investors, and a strong opinion from an influential analyst can have a significant and immediate impact upon a company's stock price. Thus, it is not surprising that there is a wide body of literature which analyzes the analysts. Prior research topics include firm characteristics and estimate accuracy[1], analyst characteristics and estimate accuracy[2], estimate dispersion and its ability to predict future volatility and returns[3],

---

[1] Das (1998), Kross *et al.* (1990), Beckers *et al.* (2004).

[2] Dugar and Nathan (1995), Clement (1997), Carelton *et al.* (1998), Desai *et al.* (2000), Hodgkinson (2001), Irvine *et al.* (2004).

[3] Bildersee *et al.* (1996), Han and Manry (2000), Athanasaakos and Kalimipalli (2003), Beckers *et al.* (2004), Johnson (2004).

estimate timeliness and accuracy[4], investors' reactions to estimates and changes therein[5], coverage initiation and suspension[6], and biases in estimates and recommendations. Of these, the last two topics will be the focus of this paper. Specifically, I attempt to determine if the dropping or suspension of coverage can explain the bias in analysts' recommendations, and if adjusting for dropped coverage can produce more accurate recommendations and estimates.

4.1.1 Dropping coverage

Analysts and the companies they cover have a strange, symbiotic relationship. Collectively, analysts have the power to cause serious damage to a company and its management, yet individually they are highly dependent upon the generosity of the company in providing access and information. Add to this the other business relations a company may have with an analyst's firm, and it becomes apparent that many analysts walk a fine line between objectivity and partiality.

Thus, it is not surprising that recommendations take on euphemistic names such as "market under perform" or "underweight" rather than "sell immediately". Similarly, it makes sense that an analyst would prefer to drop or suspend coverage rather than significantly downgrade a company and risk angering management. For example, suspension of coverage could be explained to the company's management as a temporary measure, due to a lack of resources as a result of recent and unexpected staff departures.

Given this, I hypothesize that there is information value in dropped or suspended coverage which is not being properly incorporated in the consensus recommendations and estimates of the remaining analysts. This hypothesis will be empirically examined later in the paper.

---

4 Stickel (1992), Cooper *et al.* (2001).

5 Stickel (1991), Stickel (1995), Francis and Soffer (1997), Ho and Harris (1998), Michaely and Womack (1999), Ho and Harris (2000), Krishnan and Booker (2002), Gleason (2003).

6 Bushan (1989), O'Brien and Bhushan (1990), McNichols and O'Brien (1997), Kim *et al.* (1997), Rao *et al.* (2001), Rock *et al.* (2001).

4.1.2 Biases

The majority of the literature supports the theory that analysts are overly optimistic in making their recommendations and estimates. This optimism has been attributed to several factors:

1. *Business relations*[7] - Sell-side analysts are often employed by large financial services firms, who provide both brokerage and investment banking services. In theory, strict controls are supposed to be in place to keep these divisions separated, but in reality investment banking relations (or potential relations) have exerted an influence on affiliated analysts' estimates and recommendations. (One need only look at the $1.4 billion settlement in 2004 by Citigroup and other firms for confirmation of this fact.)
2. *Self-selection*[8] - Analysts and their firms select which companies they will cover. Optimistic estimates garner more business than pessimistic estimates, as there are a virtually unlimited number of potential buyers, but only a small number of potential sellers. Additionally, the media prefers to focus on positive recommendations versus negative recommendations. Thus, it is natural that more of the companies covered will be expected to perform well.
3. *Attachment*[9] - Often, individual analysts can select the firms within an industry that they will cover. In addition to the previously mentioned self-selection bias, this can also induce an attachment bias, where the analyst becomes overly enamoured with a particular stock. This effect also occurs with equity holders who find it difficult to sell a poor performing stock for non-rational reasons. For example, they may have owned the stock for a long period or the company may produce a

[7] Dugar and Nathan (1995), Michaely and Womack (1999), Hodgkingson (2001).
[8] McNicols and O'Brien (1997), Hong (2002).
[9] Hong (2002).

product to which they have an emotional attachment, such as Walt Disney Company.

4. *Compensation* [10] - Analysts whose firms have other business relationships with the target company, such as an investment banking relationship, realize that the market assumes their recommendations and estimates will be overly optimistic. Thus, to compensate for the market's discounting, they will favourably bias their recommendations and estimates.
5. *Overreaction* [11] - Investors, and even professional analysts, often overreact to good news. For example, a small earnings increase following numerous quarters of decreased earnings is often prematurely interpreted as the beginning of a turn-around, resulting in overly optimistic recommendations and estimates.
6. *Herding*[12] - Analysts, like mutual fund managers, do not like to be too far from the average or consensus values. While the rewards from being the one rebel who is correct are significant, the punishments for being the one rebel who is wrong are even greater. Thus, analysts tend to show "an inappropriate degree of consensus in estimates relative to observed outcomes"[13], which is known as herding. When combining this behaviour with overreaction by one or two first-moving analysts, it is easy to see how a contagion effect can take hold, resulting in optimistically biased recommendations and estimates.

Increasing attention has been focused on ways to correct for these biases, with the objective of arriving at a more accurate consensus estimate. Chase (2000), drawing on prior work from decision science, looked at alternate methodologies for combining separate, but not systematically biased, estimates to achieve increased accuracy. Hayes

[10] Loffler (1998).
[11] Debondt and Thaler (1984).
[12] Debondt and Forbes (1999), Desai *et al.* (2000), Hong *et al.* (2000), Rao *et al.* (2001), Gleason and Lee (2003).
[13] Debondt (1999).

and Levine (2000) postulated that analysts' estimates are drawn from a truncated normal distribution and used maximum likelihood estimators instead of the mean value. Kim *et al.* (2001) used the mean plus a positive multiple of the change in the mean instead of the mean alone. Gu and Wu (2003) found the median estimate to be more accurate than the mean estimate.

In summary, it has been well-documented that analysts are optimistically biased in their estimates, but it has not yet been resolved what the exact origin of this bias is, nor how to correct for it. I contribute to this debate by hypothesizing that a contributing factor is *pessimistic analysts are removing themselves from the analyst pool, rather than issuing a negative recommendation.* If this hypothesis is true, then there is information content contained in the dropped or suspended coverage data that is not being incorporated in the consensus recommendations and estimates. The remainder of this paper empirically examines this hypothesis.

4.1.3 Data options

Much of the prior research, including the majority of that cited above, has been based on analysts' *estimates* of future earnings. In my analysis, I will start with a simplified analysis based upon analysts' *consensus recommendations*, and then proceed to a more detailed analysis based on analysts' *individual estimates*. My rationale for analyzing both recommendations and estimates is that each data set has relative strengths, and by analyzing both I can arrive at a more complete assessment of the impact of dropped coverage. Strengths of the recommendation data are:

1. Earnings estimates are a tool (albeit a very important one) to arrive at a recommendation, and focusing exclusively upon earnings may miss other non-earnings factors that may be of relevance (e.g. quality of earnings).
2. Earnings estimates are generally given for a maximum of three years forward. Making buy and sell decisions based solely upon these estimates may omit future events that are known with relative certainty

(e.g. mining or manufacturing facilities that will be coming on line, patents set to expire).

3. When an analyst drops a company, it is more difficult and subjective to infer an earnings estimate from this action. It is straightforward to impute a SELL[14] recommendation in such cases, rather than trying to assign a low (or negative) earnings estimate.
4. Recommendations are easily comparable across firms, whereas estimates require adjustments to account for factors such as differences in share prices and industry profitability levels.

Strengths of the estimate data are:

1. Estimates can take any penny-denominated increment, whereas recommendations come in five steps. Thus, there is less loss of detail in using estimates.[15]
2. Using detail-level data (i.e. at a *company–broker* level) allows for exact tracking of which brokers are dropping coverage, which removes potential inaccuracies introduced by approximating brokerage behaviour based on the changes in the number of analysts issuing recommendations.
3. Using estimate data allows for exact comparison with actual values for the specified interval. For example, an annual earnings per share estimate can be compared to the actual annual earnings per share value for that year. This removes the problem of mismatched time-frames, where a "long term" recommendation may be erroneously compared to the short term performance of a stock.

Full details of how the respective data sets are utilized are given in the following sections.

[14] I/B/E/S reports recommendations using a five-step system: 1=Strong Buy, 2=Buy, 3=Hold, 4=Underperform, and 5=Sell.

[15] Although the majority of analysts use some variant of the five-step rating system, some do use a three- or four-step system. I/B/E/S converts such ratings to a five-step system for reporting purposes. This can result in an even greater loss of accuracy.

4.2 CONSENSUS RECOMMENDATION ANALYSIS – UNMODIFIED RECOMMENDATIONS

4.2.1 Data

The main data set used in this section is the Institutional Brokers Estimates System (I/B/E/S) Recommendations Summary Statistics (Consensus Recommendations) data file for the date range of January 1994 to December 2006. This file has one observation per month for each company covered by one or more analyst(s), and contains consensus recommendation values based on the most recent recommendations from all analysts as of the monthly summary date.[16] I further restrict the data to U.S. companies traded on the NYSE, AMEX and NASDAQ, and I drop those consensus recommendations which have three or fewer brokers providing recommendations. This results in a final data set of 412,624 observations.

4.2.2 Distribution of recommendations

As a first step, I calculate the mean consensus recommendation for the entire data set. I/B/E/S reports recommendations using a five-step system: 1 = Strong Buy, 2 = Buy, 3 = Hold, 4 = Underperform, and 5 = Sell. Thus, were there to be no biases in the recommendations or in the selection of companies being covered, the mean should be close to 3.0 (the average of the highest rating of 1 and the lowest rating of 5). The actual mean recommendation is 2.198 with a standard deviation of 0.534. A t-test for the hypothesis that the mean recommendation is 3.0 has a t-statistic of -964.74 and a p-value of 0.000, indicating that the actual mean recommendation is different from 3.0 at all significance levels. I take this as evidence that the recommendations have an optimistic bias.

Next, I divide the mean recommendations into four ranges: between 1.0 and 2.0 (high), between 2.0 and 3.0 (high middle), between 3.0 and 4.0 (low middle) and between

[16] The summary date is typically the third Thursday of each month.

4.0 and 5.0 (low). The number of recommendations that fall into each of these ranges is shown in Figure 4.1 below.

[Figure 4.1 about here]

A very small percentage (0.2%) of recommendations fall into the low range, and the distribution is highly skewed towards the high range. Were the covered companies randomly selected from all companies listed on their respective exchange, the distributions should be more symmetric as, by definition, not all stocks can outperform the overall market.

4.2.3 Analyst performance

Next, I analyze the accuracy of analysts' consensus recommendations. This is done by calculating the average cumulative post-summary date returns by recommendation category. To do this, I use the monthly holding period returns from the CRSP Monthly data file. For each consensus recommendation observation, the cumulative monthly returns for the first through twelfth following months are calculated. These are then averaged for each of the four recommendation categories.

If analysts are providing accurate recommendations, there should be higher cumulative returns for the more positive recommendation categories, and the differences between recommendation categories should be statistically significant. To test the first requirement, I analyze the raw cumulative monthly returns by recommendation category. Results are shown in Figure 4.2 below.

[Figure 4.2 about here]

The lower rated stocks outperform the higher rated stocks in almost all time frames. In fact, for all but the first three months, returns are monotonically increasing as one goes from higher to lower rated stocks, and for the first three months this pattern is only disrupted by the returns to the low category. I have three possible explanations for this. First, analysts may be overly optimistic (pessimistic) regarding their higher (lower) recommendation stocks. Second, analysts' recommendations may be primarily based on past performance.[17] Third, analysts' recommendations may be more short-term in nature.[18] Regardless of the reason, it is clear that the recommendations are, on average, not an accurate predictor of performance in the subsequent 12 months.[19]

Next, I test the requirement that differences in returns between the rating categories be statistically significant. To do this, I perform *t*-tests for the difference between each pair of recommendation categories for each of the 12 cumulative holding period returns. The results are summarized in Table 4.1 below.

[Table 4.1 about here]

As can be seen, in the majority of cases the means are significantly different. The exceptions are those pairs which include recommendation category 4 (low). This must be interpreted with caution, though, as the high *p*-values are primarily a result of the low number of observations found in recommendation category 4 (low). For example, recommendation category 4 (low) contains approximately 600 observations in each of the cumulative return intervals, while category 3 (low middle) contains approximately 35,000 observations in each of the cumulative return intervals.

One potential limitation of my analysis is that I am using *raw* returns. This could be producing inaccurate results if certain risk characteristics are not uniformly distributed across the ratings categories, and the higher returns to the lower rated stocks are attributable to a higher level of riskiness. For example, the lower rated categories may contain a higher proportion of small stocks. To account for this, I also calculate risk-

---

17 This is in line with contrarian strategies (i.e. superior returns can be achieved by buying past losers and selling past winners) documented by researchers such as De Bondt and Thaler (1985) and Lakonkishok, *et al.* (1994).

18 Consensus recommendation summaries are typically produced by I/B/E/S on the third Thursday of every month, based on recommendations issued (or confirmed) since the last summary date. Monthly returns used in Figure 4.2 begin on the first day of the month following the consensus recommendation summary. Thus, a lag of two to six weeks may exist between the issuance of a recommendation and the inclusion of an underlying stock in a portfolio.

19 It is possible that these results are being influenced by selection bias. Specifically, stocks which have been selected for coverage may not be representative of the overall universe of stocks. To check for such bias, a Heckman correction is implemented as in Garfinkel and Sokobin (2006). There is no substantial change to the results.

adjusted returns using a three factor model.[20] Results are shown in Appendix 4.1, Panels A and B, and are qualitatively similar to those using raw returns.

Overall, I conclude that the returns to the recommendation categories are significantly different, but that the accuracy of the recommendations is highly suspect, given that negatively rated stocks outperform positively rated stocks based on both raw and adjusted returns.

4.3 CONSENSUS RECOMMENDATION ANALYSIS – MODIFIED RECOMMENDATIONS

4.3.1 Methodologies

In this section I attempt to adjust for analysts' biases by incorporating information regarding dropped coverage. The resulting distribution and accuracy of these adjusted recommendations will then be compared with those based on the unmodified results.

To make these adjustments, I utilize two methods. Under *Method 1*, a running total of the number of analysts dropping coverage (or 'drops') is retained for each company. At each summary date, the mean recommendation is recalculated by assigning a low rating (i.e. a numeric value of 5.0) to each of the outstanding drops. If the number of analysts subsequently increases, the running total of drops is increased by one for each new analyst. No adjustments are made to the mean recommendation if the running total of drops has a positive value (i.e. more analysts have initiated coverage than dropped coverage since the start of the data set), but the running total of drops is still retained and carried forward.

The rationale behind this method is that when an analyst drops coverage, it is equivalent to issuing a low recommendation (i.e. a numeric value of 5.0) which remains in effect until coverage is re-initiated. When the same (or another) analyst subsequently

[20] Risk-adjusted returns are the residuals from monthly regressions (by company) of actual returns less risk-free returns on the typical Fama-French three factors (market return less risk-free return, small minus big (size) returns and high minus low (B/M) returns). The monthly Fama-French factors are sourced from Ken French's data library available through his website.

initiates coverage[21], it is no longer assumed that they are issuing a low recommendation. Instead, their recommendation is now incorporated into the reported consensus recommendation value. Thus, positive values for the running total of analyst changes do not result in modifications to the mean recommendation, as the analysts represented by that positive number already have their recommendations included in the mean recommendation value.

Under *Method 2*, a ceiling of zero is imposed on the running total of analyst changes. Thus, an addition followed by a drop will result in a running total of -1, as compared to 0 under *Method 1*. By placing a ceiling of zero on the cumulative number of analyst changes, I control for the situation where a newly-listed company develops an analyst following during the data period and is incorrectly credited with a large positive "buffer" simply for adjusting to a normal level of coverage.

Numeric examples of both methods are included in Appendix 4.2.

4.3.2 Distribution of recommendations with bias-correcting modifications

Mean recommendations are recalculated using the above two methods, and are allocated to the same four ranges previously used. Results are shown in Figure 4.3 below, along with the unmodified results from the prior section.

[Figure 4.3 about here]

Method 1 only slightly reduces the skewness in the distribution. By contrast, *Method 2* provides a significant shift towards a more normal distribution.

I next calculate the overall mean recommendations, and run *t*-tests run to see if they are statistically different than 3.0. Results for the two modification methods, along with a repeat of the unmodified results, are shown in Table 4.2 below.

[21] With the summary recommendation data set, it cannot be determined if an analyst who previously dropped coverage is re-initiating coverage, or if an entirely new analyst is initiating coverage. For the calculations in this section the distinction is irrelevant.

[Table 4.2 about here]

Methods 1 and *2* move the mean recommendation progressively closer to the desired mean value of 3.0, although the *p*-values indicate that a statistically significant level of bias still exists.

4.3.3 Analyst performance with bias-correcting modifications

Next, I repeat the prior analysis of the accuracy of analysts' recommendations using the modified recommendations. Results for *Method 1* are shown in Figure 4.4 and Table 4.3 below, and results for *Method 2* are shown in Figure 4.5 and Table 4.4 below.

[Figure 4.4 about here]

[Table 4.3 about here]

[Figure 4.5 about here]

[Table 4.4 about here]

The results are very similar to those based on the unmodified recommendation categories. For both *Methods 1* and *2*, the returns increase monotonically for all month ranges when going from high recommendations to low recommendations. Also, *p*-values for differences between category pairs have improved, and in the case of *Method 2* all pairs are significantly different. Even so, the returns are still the opposite of what would be expected if the recommendations are accurate. Specifically, stocks with lower recommendations still outperform those with higher recommendations.

These analyses are also repeated using the three-factor adjusted returns as described in the prior section. The results are shown in Appendix 4.1, Panels C through F, and are qualitatively similar to those using raw returns.

Overall, I conclude that while the modification methods, especially *Method 2*, do result in a more normal distribution of recommendations, they are unable to improve the accuracy of the recommendations.

4.4 DETAILED ESTIMATE ANALYSIS

4.4.1 Introduction

In the prior section I utilize I/B/E/S summary recommendation data, and find that replacing dropped recommendations with the lowest recommendation value reduces the skewness of the recommendation distribution, but does not significantly improve the accuracy of the consensus recommendations.

Unfortunately, this analysis has several limitations. First, it uses a five-category rating system, which results in a loss of detail when I/B/E/S is required to translate from a different rating system. [22] Second, there could be significant differences in recommendations within each category. Third, the analysis looks at one to 12 month performance, while the recommendation may be more long-term in nature. Fourth, it does not track specific brokers, but rather makes assumptions regarding which ones drop and re-initiate coverage.

To overcome these limitations, I implement an analysis which utilizes the I/B/E/S Detail History data file. This file contains estimated and actual annual earnings per share (EPS) values by broker, which will avoid the problems noted above. By using exact EPS estimates and actual values, problems one and two regarding the potential over-aggregation and/or misclassification of recommendation categories are removed. Similarly, by comparing estimates and actual values for a specific fiscal year, the issue noted in problem three is removed, as the estimates and actual values are for the exact same time period. Also, the fourth problem is avoided by using broker-level estimates, as values can be substituted when a specific broker drops coverage, and removed if and when that broker resumes coverage.

---

[22] For example, some brokers use three or four recommendation categories.

4.4.2 Data and analysis

I begin with the full I/B/E/S Detail History data file. From this, I select only EPS estimates for the current fiscal year made between 1994 and 2006, inclusive. The data set is further filtered to include only U.S. firms[23] traded on the NYSE, AMEX and NASDAQ.[24] This results in 420,507 estimates across 15,584 companies and 847 brokers.

Next, for each company-year, the following summary statistics are calculated: number of estimates, mean estimate, and standard deviation of estimates. The mean number of estimates per company-year is 7.49, with a standard deviation of 7.51 and a range of 1 to 56.

Each series of annual estimates for a company-broker is then examined to determine situations where coverage has been dropped.[25] A broker is considered to have dropped coverage if it provides estimates for one or more fiscal years, and then stops providing estimates for one or more fiscal years. If coverage subsequently resumes, it is only the intervening missing years that will be treated as dropped years.

When substituting values for dropped coverage, there are two considerations. The first is the number of years for which to substitute values. To evaluate the effectiveness of different methodologies, I run the subsequent analyses using substitution periods of 1, 2, 3, 4, 5 and an unlimited number of years.[26] The second consideration is the specific substitution value to be used. Again, I implement several different methods:

- *Method A:* Estimate is set to the minimum of remaining estimates
- *Method B:* Estimate is set to the mean of remaining estimates less ½ times the standard deviation of remaining estimates (i.e. mean – 0.5 × std dev)

---

23 I/B/E/S selection variables used are MEASURE = "EPS" and FPI = "1".

24 Where a broker issues more than one estimate for a given company / fiscal year, I select only the most recent estimate prior to the earnings announcement date.

25 Note that I/B/E/S also provides analyst information, which is a further level of detail beyond the broker level. I do not utilize this data, as I am not concerned with which analyst within a given brokerage is covering a stock, but rather whether or not the brokerage is covering a stock at all. This avoids issues where specific analysts change employers, where coverage responsibilities are shifted among analysts within a brokerage, or where analyst information is not reported by the brokerage.

26 Note that if another estimate is made before the end of the substitution period, the substitution period ends and that estimate is used. Similarly, no substitution periods are allowed to extend the date range beyond 2006.

- *Method C*: Estimate is set to the mean of remaining estimates less 1 times the standard deviation of remaining estimates (i.e. mean – 1.0 × std dev)
- *Method D:* Estimate is set to the mean of remaining estimates less 1½ times the standard deviation of remaining estimates (i.e. mean – 1.5 × std dev)
- *Method E:* Estimate is set to the mean of remaining estimates less 2 times the standard deviation of remaining estimates (i.e. mean – 2.0 × std dev)

This results in a total of 30 different analyses, which vary in the number of estimates being filled in, and the magnitude of the substituted values.

For each of these combinations of substitution years and calculation method, I calculate the mean EPS estimate by company-year. I then calculate a price-scaled estimate delta value as the mean EPS estimate minus the actual EPS value, all divided by the prior year-end closing price.[27] Next, the price-scaled delta values are averaged across all company-years, and *p*-values are calculated for the hypothesis that the mean price-scaled delta value is equal to zero (i.e. perfect estimate accuracy). Results are shown in Table 4.5 below.

[Table 4.5 about here]

There are two main observations to be made about the results. First, in terms of the magnitude of the mean price-scaled deltas, all of the modified estimation methods produce better results than the unmodified base case. Second, based on *p*-values, only two Method–year combinations have a significantly high *p*-value: *Method A* with a maximum of three substitution years, and *Method D* with a maximum of two substitution

[27] Two filtering criteria are applied at this point. First, observations for company-years with three or fewer (non-substituted) estimates are removed. Second, price-scaled delta values are trimmed at the 1^{st} / 99^{th} percentiles calculated across all remaining observations.

years. (Recall that *Method A* uses the minimum remaining estimate as the substitution value, and *Method D* uses the mean estimate less 1.5 times the standard deviation of estimates as the substitution value.)

Investigating these two results further reveals that the minimum remaining estimate is on average -1.33 standard deviations away from the mean. This makes sense, as the substituted estimates in *Method A* are slightly smaller in magnitude than those in *Method D*, and thus give similar results when applied for one extra year. Next, I try to improve on both the *Method A* – 3 year and the *Method D* – 2 year results.

Given that I am using annual estimates, there are no further adjustments that can be made based on substitution years. For example, *Method A – 2 year* and *Method A – 4 year* results (i.e. using +/- one year) have already been generated and are found to have insignificant *p*-values. On the other hand, the substitution values can be adjusted. For both the *Method A – 3 year* and the *Method D – 2 year* results I re-run the analyses with adjustments of +/-0.01 standard deviations. For the *3 year* results I start at -1.33 standard deviations from the mean, and for the *2 year* results I start at -1.50 standard deviations from the mean. Following an iterative procedure, I find that the following standard deviation–year combinations, shown in Table 4.6 below, give the lowest deltas and highest *p*-values.

[Table 4.6 about here]

Using two substitution years, the substitution value of *mean estimate – 1.50 std dev* proves to be the most accurate value. (For example, using a substitution value based on *1.49 (1.51) standard deviations* results in a lower *p*-value of 0.9181 (0.8079)). Using three substitution years, a substitution value of *mean estimate – 1.27× std dev* results in the most accurate value. Between the two, the ***3 year, -1.27 std dev*** method is slightly superior as it has a lower mean delta and standard deviation, and a higher *p*-value.

I interpret these results as follows. When an analyst drops coverage, it can be interpreted as issuing an EPS estimate of either (a) the mean remaining estimate less 1.5 standard deviations of the remaining estimates for the following two years, or (b) the mean estimate less 1.27 standard deviations of the remaining estimates for the following three years. Alternately, a slightly less accurate "rule of thumb" is to assume dropped coverage is analogous issuing an estimate equal to the minimum of the remaining estimates for the following three years.

A potential limitation of this analysis is that it is based on annual estimates. As previously mentioned, this restricts adjustments to the substitution period to whole year increments, and effectively makes the substitution value the main variable by which the methodologies can be adjusted. A possible future expansion would be to use quarterly estimates, which would allow for a greater degree of adjustment to the substitution periods.

An additional expansion involves more specific controls for dropped coverage. For example, by analyzing the overall number of analysts providing coverage for the entire universe of stocks, dropped coverage due to economic downturns and financial industry lay-offs could be controlled for, allowing for isolation of those situations where coverage is dropped due to firm-specific factors.

4.5 CONCLUSION

Stock market analysts are themselves highly analyzed, and a large body of literature exists which evaluates their behaviour and predictions. Even so, the majority of this literature has focused on earnings estimates, not overall recommendations. Similarly, very little investigation has been performed on the impact of dropping or suspending coverage. I address both of these shortcomings by evaluating the impact of dropped coverage on the bias and accuracy of analysts' consensus recommendations, and on the accuracy of analysts' annual earnings per share estimates.

As a result, this paper contributes to the existing literature in three ways. First, it provides additional evidence that analysts' recommendations are optimistically biased, and poorly correlated with subsequent returns. Second, it shows that incorporating the information available from dropped coverage can significantly reduce the optimism bias in analysts' recommendations. Third, this incorporation does not significantly improve the accuracy of the recommendations, but it can significantly improve the accuracy of the annual earnings per share estimates. This is taken as evidence that there is significant information content in the number of analysts which drop or suspend coverage for a given company.

Table 4.1 – P-values for comparison of cumulative raw return means by rating category pairs

	1 vs. 2	1 vs. 3	1 vs. 4	2 vs. 3	2 vs. 4	3 vs. 4
1 month	0.000	0.000	0.709	0.003	0.372	0.220
2 months	0.000	0.000	0.574	0.000	0.796	0.380
3 months	0.000	0.000	0.108	0.000	0.606	0.746
4 months	0.000	0.000	0.010	0.000	0.212	0.812
5 months	0.000	0.000	0.000	0.000	0.039	0.363
6 months	0.000	0.000	0.000	0.000	0.005	0.129
7 months	0.000	0.000	0.000	0.000	0.002	0.081
8 months	0.000	0.000	0.000	0.000	0.000	0.028
9 months	0.000	0.000	0.000	0.000	0.000	0.015
10 months	0.000	0.000	0.000	0.000	0.000	0.012
11 months	0.000	0.000	0.000	0.000	0.000	0.026
12 months	0.000	0.000	0.000	0.000	0.000	0.017

Rating categories: 1 = high, 2 = high middle, 3 = low middle, 4 = low.

Table 4.2– Summary statistics and *t*-test results for H0:mean recommendation = 3.0

	Mean	Std Dev	t-stat	p-value
Unadjusted	2.198	0.534	-964.744	0.000
Method 1	2.263	0.594	-796.498	0.000
Method 2	2.611	0.738	-338.952	0.000

Table 4.3 – P-values for comparison of cumulative raw return means by rating category pairs using modification method 1

	1 vs. 2	1 vs. 3	1 vs. 4	2 vs. 3	2 vs. 4	3 vs. 4
1 month	0.000	0.000	0.045	0.003	0.435	0.895
2 months	0.000	0.000	0.000	0.000	0.110	0.633
3 months	0.000	0.000	0.000	0.000	0.011	0.273
4 months	0.000	0.000	0.000	0.000	0.002	0.178
5 months	0.000	0.000	0.000	0.000	0.000	0.119
6 months	0.000	0.000	0.000	0.000	0.000	0.056
7 months	0.000	0.000	0.000	0.000	0.000	0.018
8 months	0.000	0.000	0.000	0.000	0.000	0.002
9 months	0.000	0.000	0.000	0.000	0.000	0.000
10 months	0.000	0.000	0.000	0.000	0.000	0.000
11 months	0.000	0.000	0.000	0.000	0.000	0.000
12 months	0.000	0.000	0.000	0.000	0.000	0.000

Rating categories: 1 = high, 2 = high middle, 3 = low middle, 4 = low.

Table 4.4 – P-values for comparison of cumulative raw return means by rating category pairs using modification method 2

	1 vs. 2	1 vs. 3	1 vs. 4	2 vs. 3	2 vs. 4	3 vs. 4
1 month	0.000	0.000	0.000	0.000	0.000	0.000
2 months	0.000	0.000	0.000	0.000	0.000	0.000
3 months	0.000	0.000	0.000	0.000	0.000	0.000
4 months	0.000	0.000	0.000	0.000	0.000	0.000
5 months	0.000	0.000	0.000	0.000	0.000	0.000
6 months	0.000	0.000	0.000	0.000	0.000	0.000
7 months	0.000	0.000	0.000	0.000	0.000	0.000
8 months	0.000	0.000	0.000	0.000	0.000	0.000
9 months	0.000	0.000	0.000	0.000	0.000	0.000
10 months	0.000	0.000	0.000	0.000	0.000	0.000
11 months	0.000	0.000	0.000	0.000	0.000	0.000
12 months	0.000	0.000	0.000	0.000	0.000	0.000

Rating categories: 1 = high, 2 = high middle, 3 = low middle, 4 = low.

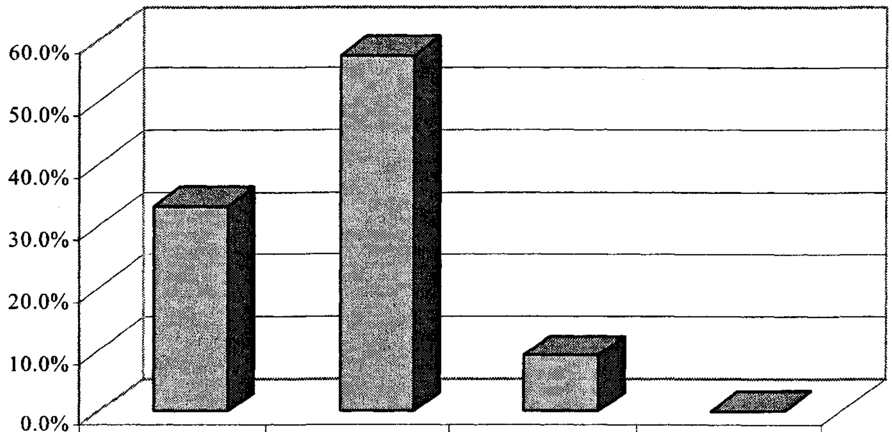
Table 4.5 – EPS estimate accuracy by substitution years and calculation method

Substitution Years	Mean of Delta	Std Error	Min Delta	Max Delta	*t*-stat H0:Mean Delta=0	*p*-value H0:Mean Delta=0
BASE (Unmodified) CASE						
N/A	0.00259	0.00009	-0.0557	0.1680	28.6402	0.0000
METHOD A						
1	0.00111	0.00010	-1.2884	0.1675	11.3269	0.0000
2	0.00040	0.00010	-1.4674	0.1675	3.8982	0.0001
3	-0.00002	0.00011	-1.5415	0.1675	-0.1453	0.8844
4	-0.00025	0.00011	-1.5595	0.1675	-2.2851	0.0223
5	-0.00038	0.00011	-1.5652	0.1675	-3.4433	0.0006
Unlimited	-0.00053	0.00011	-1.5707	0.1675	-4.8076	0.0000
METHOD B						
1	0.00200	0.00009	-0.4211	0.1675	22.7501	0.0000
2	0.00173	0.00009	-0.4796	0.1675	19.8101	0.0000
3	0.00157	0.00009	-0.5039	0.1675	18.0980	0.0000
4	0.00149	0.00009	-0.5098	0.1675	17.1612	0.0000
5	0.00145	0.00009	-0.5116	0.1675	16.6478	0.0000
Unlimited	0.00140	0.00009	-0.5134	0.1675	16.0839	0.0000
METHOD C						
1	0.00141	0.00009	-0.8423	0.1675	15.7165	0.0000
2	0.00086	0.00009	-0.9593	0.1675	9.4907	0.0000
3	0.00055	0.00009	-1.0077	0.1675	6.0211	0.0000
4	0.00039	0.00009	-1.0195	0.1675	4.1709	0.0000
5	0.00030	0.00009	-1.0232	0.1675	3.1933	0.0014
Unlimited	0.00020	0.00009	-1.0268	0.1675	2.1238	0.0337
METHOD D						
1	0.00081	0.00010	-1.2634	0.1675	8.5494	0.0000
2	-0.00001	0.00010	-1.4389	0.1675	-0.0706	0.9437
3	-0.00047	0.00010	-1.5116	0.1675	-4.4938	0.0000
4	-0.00072	0.00011	-1.5293	0.1675	-6.7603	0.0000
5	-0.00085	0.00011	-1.5348	0.1675	-7.9319	0.0000
Unlimited	-0.00100	0.00011	-1.5402	0.1675	-9.2308	0.0000
METHOD E						
1	0.00022	0.00010	-1.6846	0.1675	2.1073	0.0351
2	-0.00087	0.00011	-1.9185	0.1675	-7.6448	0.0000
3	-0.00149	0.00012	-2.0154	0.1675	-12.2575	0.0000
4	-0.00182	0.00013	-2.0390	0.1675	-14.5468	0.0000
5	-0.00200	0.00013	-2.0464	0.1675	-15.7244	0.0000
Unlimited	-0.00220	0.00013	-2.0536	0.1675	-17.0607	0.0000

Table 4.6 – Optimal EPS substitution values

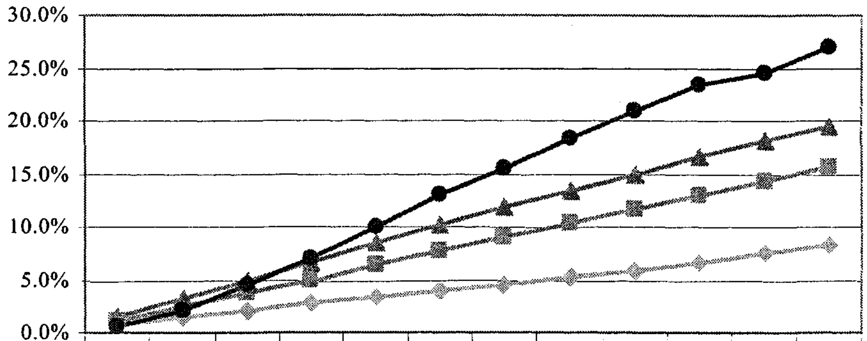
Substitution Years	Substitution Value	Mean Delta	Std Dev Delta	Min Delta	Max Delta	p-value H0:Mean Delta=0
2	Mean - 1.50 Std Dev	-0.000007	0.0183	-1.4389	0.1675	0.9437
3	Mean - 1.27 Std Dev	0.000002	0.0179	-1.2798	0.1675	0.9849

Figure 4.1 – Percent of mean recommendations by range



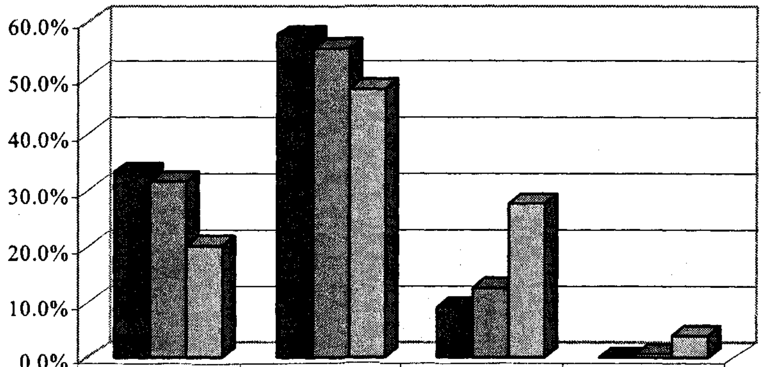
	1.0 to 2.0 (high)	2.0 to 3.0 (high middle)	3.0 to 4.0 (low middle)	4.0 to 5.0 (low)
Percent of Total Obs	33.2%	57.6%	9.0%	0.2%

Figure 4.2 – Cumulative monthly raw returns by mean recommendation category



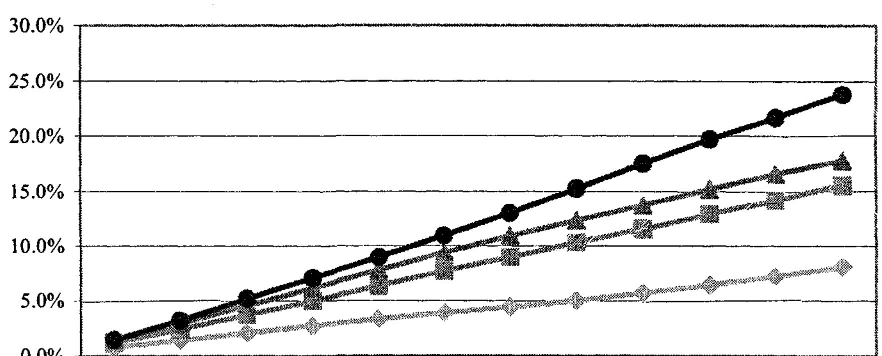
	1	2	3	4	5	6	7	8	9	10	11	12
1.0 to 2.0 (high)	0.8%	1.5%	2.1%	2.8%	3.4%	4.0%	4.6%	5.2%	5.9%	6.7%	7.5%	8.3%
2.0 to 3.0 (high middle)	1.2%	2.4%	3.7%	5.0%	6.3%	7.7%	9.0%	10.3%	11.6%	13.0%	14.3%	15.6%
3.0 to 4.0 (low middle)	1.5%	3.1%	4.9%	6.6%	8.4%	10.2%	11.8%	13.4%	14.9%	16.6%	18.1%	19.4%
4.0 to 5.0 (low)	0.5%	2.1%	4.4%	7.0%	10.1%	13.1%	15.6%	18.4%	20.9%	23.3%	24.5%	27.1%

Figure 4.3 – Percent of mean recommendations by range by modification method



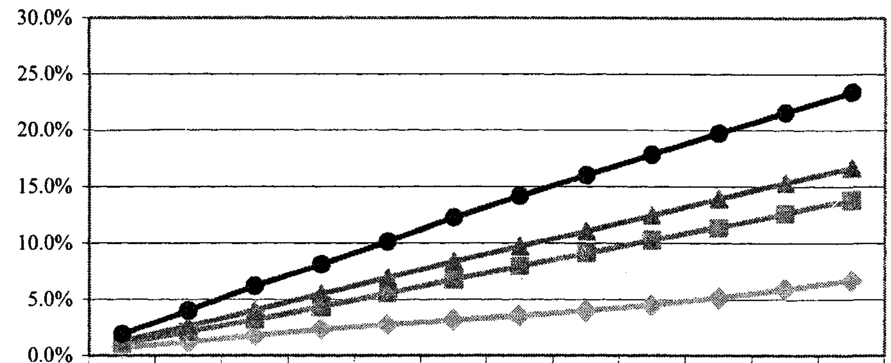
	1.0 to 2.0 (high)	2.0 to 3.0 (high middle)	3.0 to 4.0 (low middle)	4.0 to 5.0 (low)
Unadjusted	33.2%	57.6%	9.0%	0.2%
Method 1	31.6%	55.3%	12.5%	0.7%
Method 2	20.2%	48.0%	27.8%	4.0%

Figure 4.4 – Cumulative monthly raw returns by mean recommendation category using modification method 1



	1	2	3	4	5	6	7	8	9	10	11	12
1.0 to 2.0 (high)	0.8%	1.4%	2.1%	2.7%	3.3%	3.9%	4.4%	5.0%	5.7%	6.4%	7.2%	8.1%
2.0 to 3.0 (high middle)	1.2%	2.4%	3.7%	5.0%	6.3%	7.7%	9.0%	10.3%	11.6%	12.9%	14.2%	15.6%
3.0 to 4.0 (low middle)	1.4%	2.9%	4.5%	6.1%	7.8%	9.4%	10.9%	12.3%	13.7%	15.2%	16.6%	17.8%
4.0 to 5.0 (low)	1.4%	3.2%	5.2%	7.1%	9.0%	10.9%	13.0%	15.2%	17.5%	19.7%	21.6%	23.8%

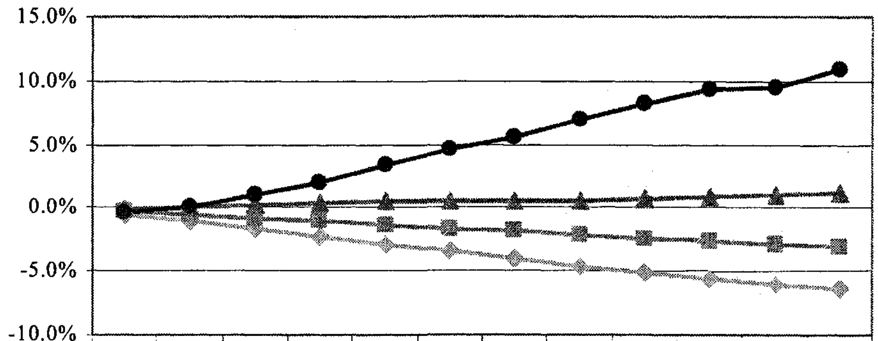
Figure 4.5 – Cumulative monthly raw returns by mean recommendation category using modification method 2



	1	2	3	4	5	6	7	8	9	10	11	12
1.0 to 2.0 (high)	0.7%	1.2%	1.8%	2.4%	2.8%	3.2%	3.6%	4.1%	4.6%	5.2%	5.9%	6.7%
2.0 to 3.0 (high middle)	1.1%	2.1%	3.2%	4.4%	5.6%	6.8%	8.0%	9.1%	10.3%	11.5%	12.6%	13.8%
3.0 to 4.0 (low middle)	1.3%	2.7%	4.1%	5.5%	6.9%	8.4%	9.7%	11.1%	12.5%	13.9%	15.3%	16.7%
4.0 to 5.0 (low)	1.9%	4.0%	6.2%	8.1%	10.2%	12.3%	14.2%	16.0%	17.8%	19.7%	21.6%	23.4%

APPENDIX 4.1 – Analyst Performance Using Three-Factor Adjusted Returns

Panel A – Cumulative Monthly Adjusted Returns by Mean Recommendation Category (Unmodified Consensus Recommendation)



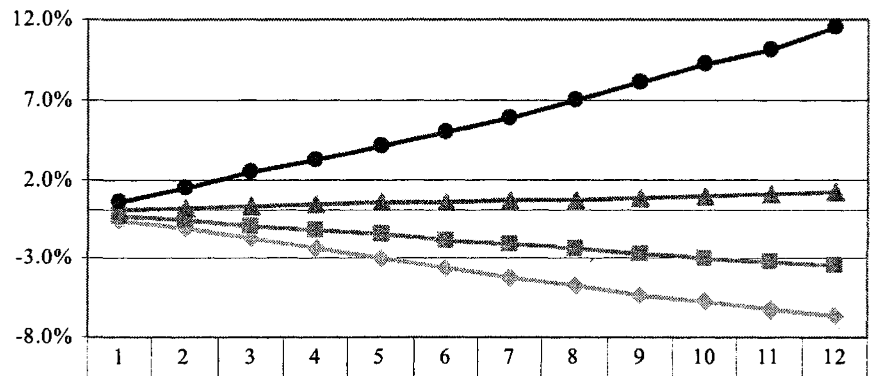
	1	2	3	4	5	6	7	8	9	10	11	12
1.0 to 2.0 (high)	-0.5%	-1.1%	-1.7%	-2.3%	-2.9%	-3.5%	-4.1%	-4.6%	-5.1%	-5.6%	-6.0%	-6.5%
2.0 to 3.0 (high middle)	-0.3%	-0.6%	-0.8%	-1.1%	-1.4%	-1.6%	-1.9%	-2.1%	-2.4%	-2.6%	-2.9%	-3.1%
3.0 to 4.0 (low middle)	0.0%	0.1%	0.3%	0.4%	0.5%	0.5%	0.5%	0.5%	0.6%	0.9%	1.0%	1.1%
4.0 to 5.0 (low)	-0.4%	0.1%	1.1%	2.0%	3.4%	4.7%	5.6%	7.0%	8.2%	9.3%	9.4%	10.9%

Panel B – P-values for Comparison of Cumulative Adjusted Return Means by Rating Category Pairs (Unmodified Consensus Recommendation)

	1 vs. 2	**1 vs. 3**	**1 vs. 4**	**2 vs. 3**	**2 vs. 4**	**3 vs. 4**
1 month	0.000	0.000	0.912	0.000	0.848	0.572
2 months	0.000	0.000	0.256	0.000	0.518	0.993
3 months	0.000	0.000	0.035	0.000	0.149	0.554
4 months	0.000	0.000	0.004	0.000	0.039	0.287
5 months	0.000	0.000	0.000	0.000	0.004	0.076
6 months	0.000	0.000	0.000	0.000	0.000	0.017
7 months	0.000	0.000	0.000	0.000	0.000	0.009
8 months	0.000	0.000	0.000	0.000	0.000	0.002
9 months	0.000	0.000	0.000	0.000	0.000	0.001
10 months	0.000	0.000	0.000	0.000	0.000	0.000
11 months	0.000	0.000	0.000	0.000	0.000	0.001
12 months	0.000	0.000	0.000	0.000	0.000	0.001

Rating categories: 1 = high, 2 = high middle, 3 = low middle, 4 = low.

Panel C – Cumulative Monthly Adjusted Returns by Mean Recommendation Category (Method 1 Modifications)



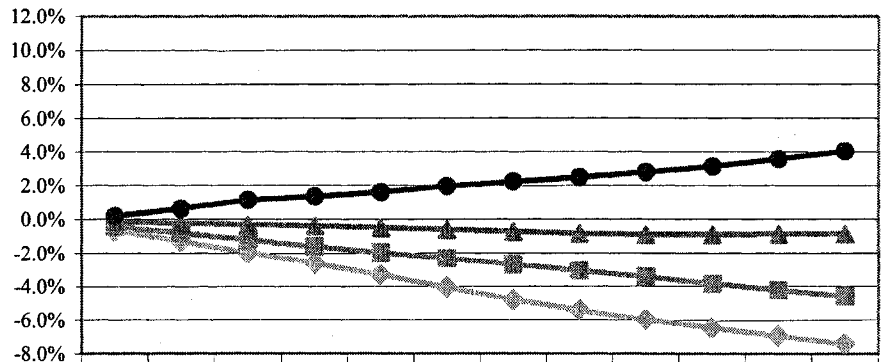
	1	2	3	4	5	6	7	8	9	10	11	12
1.0 to 2.0 (high)	-0.5%	-1.1%	-1.8%	-2.4%	-3.0%	-3.6%	-4.2%	-4.8%	-5.3%	-5.8%	-6.2%	-6.7%
2.0 to 3.0 (high middle)	-0.3%	-0.6%	-0.9%	-1.2%	-1.5%	-1.8%	-2.1%	-2.4%	-2.7%	-3.0%	-3.2%	-3.4%
3.0 to 4.0 (low middle)	0.0%	0.1%	0.3%	0.4%	0.5%	0.6%	0.7%	0.7%	0.8%	0.9%	1.1%	1.1%
4.0 to 5.0 (low)	0.6%	1.4%	2.4%	3.3%	4.1%	4.9%	5.8%	7.0%	8.1%	9.2%	10.2%	11.5%

Panel D – P-values for Comparison of Cumulative Adjusted Return Means by Rating Category Pairs (Method 1 Modifications)

	1 vs. 2	**1 vs. 3**	**1 vs. 4**	**2 vs. 3**	**2 vs. 4**	**3 vs. 4**
1 month	0.000	0.000	0.000	0.000	0.003	0.068
2 months	0.000	0.000	0.000	0.000	0.000	0.005
3 months	0.000	0.000	0.000	0.000	0.000	0.000
4 months	0.000	0.000	0.000	0.000	0.000	0.000
5 months	0.000	0.000	0.000	0.000	0.000	0.000
6 months	0.000	0.000	0.000	0.000	0.000	0.000
7 months	0.000	0.000	0.000	0.000	0.000	0.000
8 months	0.000	0.000	0.000	0.000	0.000	0.000
9 months	0.000	0.000	0.000	0.000	0.000	0.000
10 months	0.000	0.000	0.000	0.000	0.000	0.000
11 months	0.000	0.000	0.000	0.000	0.000	0.000
12 months	0.000	0.000	0.000	0.000	0.000	0.000

Rating categories: 1 = high, 2 = high middle, 3 = low middle, 4 = low.

Panel E – Cumulative Monthly Adjusted Returns by Mean Recommendation Category (Method 2 Modifications)



	1	2	3	4	5	6	7	8	9	10	11	12
1.0 to 2.0 (high)	-0.6%	-1.3%	-1.9%	-2.6%	-3.3%	-4.0%	-4.8%	-5.4%	-5.9%	-6.5%	-6.9%	-7.4%
2.0 to 3.0 (high middle)	-0.4%	-0.8%	-1.2%	-1.6%	-2.0%	-2.3%	-2.7%	-3.0%	-3.4%	-3.8%	-4.2%	-4.6%
3.0 to 4.0 (low middle)	-0.1%	-0.2%	-0.3%	-0.4%	-0.5%	-0.6%	-0.7%	-0.9%	-0.9%	-0.9%	-0.9%	-0.9%
4.0 to 5.0 (low)	0.2%	0.6%	1.1%	1.4%	1.6%	2.0%	2.2%	2.5%	2.8%	3.1%	3.6%	4.0%

Panel F – P-values for Comparison of Cumulative Adjusted Return Means by Rating Category Pairs (Method 2 Modifications)

	1 vs. 2	1 vs. 3	1 vs. 4	2 vs. 3	2 vs. 4	3 vs. 4
1 month	0.000	0.000	0.000	0.000	0.000	0.015
2 months	0.000	0.000	0.000	0.000	0.000	0.000
3 months	0.000	0.000	0.000	0.000	0.000	0.000
4 months	0.000	0.000	0.000	0.000	0.000	0.000
5 months	0.000	0.000	0.000	0.000	0.000	0.000
6 months	0.000	0.000	0.000	0.000	0.000	0.000
7 months	0.000	0.000	0.000	0.000	0.000	0.000
8 months	0.000	0.000	0.000	0.000	0.000	0.000
9 months	0.000	0.000	0.000	0.000	0.000	0.000
10 months	0.000	0.000	0.000	0.000	0.000	0.000
11 months	0.000	0.000	0.000	0.000	0.000	0.000
12 months	0.000	0.000	0.000	0.000	0.000	0.000

Rating categories: 1 = high, 2 = high middle, 3 = low middle, 4 = low.

APPENDIX 4.2 – Sample Bias Adjustment Calculations for Consensus Recommendations

	From I/B/E/S Data		Calculated							
	A	B	C	D	E	F	G	H	I	J
Summary Date	Number of Analysts	Mean Rec.	Change in Analysts	Cummulative Change in Analysts	Analyst "Drops" to Use	Adjustment to Mean (E x 5)	Analysts x Mean Rec. (A x B)	New Rec. Total (F + G)	New Analysts Total (A + E)	New Mean Rec. (H / I)
EXAMPLE 1 - Normal Fluctuations in Coverage										
Method 1										
Nov-93	10	3.00	0	0	0	0	30.00	30.00	10	3.00
Dec-93	8	3.25	-2	-2	2	10	26.00	36.00	10	3.60
Jan-94	9	3.10	1	-1	1	5	27.90	32.90	10	3.29
Feb-94	12	2.75	3	***2***	0	0	33.00	33.00	12	2.75
Mar-94	11	2.75	-1	***1***	***0***	***0***	30.25	30.25	***11***	***2.75***
Method 2										
Nov-93	10	3.00	0	0	0	0	30.00	30.00	10	3.00
Dec-93	8	3.25	-2	-2	2	10	26.00	36.00	10	3.60
Jan-94	9	3.10	1	-1	1	5	27.90	32.90	10	3.29
Feb-94	12	2.75	3	0	0	0	33.00	33.00	12	2.75
Mar-94	11	2.75	-1	-1	1	5	30.25	35.25	12	2.94
EXAMPLE 2 - Newly Added Stock with Subsequent Drop in Coverage										
Method 1										
Nov-93	1	3.00	0	0	0	0	3.00	3.00	1	3.00
Dec-93	4	3.25	3	3	0	0	13.00	13.00	4	3.25
Jan-94	10	3.10	6	9	0	0	31.00	31.00	10	3.10
Feb-94	9	2.75	-1	***8***	***0***	***0***	24.75	***24.75***	***9***	***2.75***
Mar-94	8	2.75	-1	***7***	***0***	***0***	22.00	***22.00***	***8***	***2.75***
Method 2										
Nov-93	1	3.00	0	0	0	0	3.00	3.00	1	3.00
Dec-93	4	3.25	3	0	0	0	13.00	13.00	4	3.25
Jan-94	10	3.10	6	0	0	0	31.00	31.00	10	3.10
Feb-94	9	2.75	-1	-1	1	5	24.75	29.75	10	2.98
Mar-94	8	2.75	-1	-1	2	10	22.00	32.00	10	3.20

NOTE: ***Bold*** values highlight the differences between Methods 1 and 2

BIBLIOGRAPHY

Athanassakos, George and Madhu Kalimipalli, 2003, Analyst Forecast Dispersion and Future Stock Return Volatility, *Quarterly Journal of Business and Economics* 42, 57-78.

Beckers, Stan, Michael Steliaros and Alexander Thompson, 2000, Bias in European Analysts' Earnings Forecasts, *Financial Analysts Journal* 60, 74-85.

Bildersee, John, Suresh Radhakkrishnan and Joshua Ronen, 1996, Dispersion of Analysts' Forecasts, Precision of Earnings and Trading Volume, *International Review of Financial Analysis* 5, 99-111.

Bhushan, Ravi, 1989, Firm Characteristics and Analyst Following, *Journal of Accounting and Economics* 11, 255-274.

Carelton, Willard, Carl Chen and Thomas Steiner, 1998, Optimism Biases Among Brokerage and Non-Brokerage Firms' Equity Recommendations: Agency Costs in the Investment Industry, *Financial Management* 27, 17-30.

Chase, Charles, 2000, Composite Forecasting: Combining Forecasts for Improved Accuracy, *The Journal of Business Forecasting Methods & Systems* 19, 19-22.

Clement, Michael, 1997, Analyst Forecast Accuracy: Do Ability, Resources and Portfolio Complexity Matter?, *Journal of Accounting and Economics* 27, 285-303.

Cooper, Rick, Theodore Day and Craig Lewis, 2001, Following the Leader: A Study of Individual Analysts' Earnings Forecasts, *Journal of Financial Economics* 61, 383–416

Das, Somnath, 1999, Financial Analysts' Earnings Forecasts for Loss Firms, *Managerial Finance* 24, 37-46.

De Bondt, Werner and Richard Thaler, 1984, Does the Stock Market Overreact?, *The Journal of Finance* 40, 793-805.

De Bondt, Werner and William Forbes, 1999, Herding in Analyst Earnings Forecasts: Evidence From the United Kingdom, *European Financial Management* 5, 143-163.

Desai, Hemang, Bing Liang and Ajai Singh, 2000, Do All-Stars Shine? Evaluation of Analyst Recommendations, *Financial Analysts Journal* May-June 2000, 20-29.

Dugar, Amitabh and Siva Nathan, 1995, The Effect of Investment Banking Relationships on Financial Analysts' Earnings Forecasts and Investment Recommendations, *Contemporary Accounting Research* 12, 131-160.

Francis, Jennifer and Leonard Soffer, 1997, The Relative Informativeness of Analysts' Stock Recommendations and Earnings Forecast Revisions, *Journal of Accounting Research* 35, 193-211.

Garfinkel, Jon A. and Jonathan Sokobin, 2006, Volume, Opinion Divergence, and Returns: A Study of Post-Earnings Announcement Drift, *Journal of Accounting Research* 44, 85-112.

Gleason, Cristi and Charles Lee, 2003, Analyst Forecast Revisions and Market Price Discovery, *The Accounting Review* 78, 193-225.

Gu, Zhaoyang and Joanna Wu, 2003, Earnings Skewness and Analyst Forecast Bias, *Journal of Accounting and Economics* 35, 5-29.

Han, Bong and David Manry, 2000, The Implications of Dispersion in Analysts' Earnings Forecasts for Future ROE and Future Returns, *Journal of Business Finance & Accounting*, 27, 99-125.

Hayes, Rachel and Carolyn Levine, 2000, An Approach to Adjusting Analysts' Consensus Forecasts for Selection Bias, *Contemporary Accounting Research* 17, 61-83.

Ho, Michael and Robert Harris, 1998, Market Reactions to Messages from Brokerage Ratings Systems, *Financial Analysts Journal* 54, 49-57.

Ho, Michael and Robert Harris, 2000, Brokerage Analysts' Rationale for Investment Recommendations: Market Responses to Different Types of Information, *Journal of Financial Research* 23, 449-468.

Hodgkinson, Lynn, 2001, Analysts' Forecasts and the Broker Relationship, *Journal of Business Finance & Accounting* 28, 943-961.

Hong, Harrison, Jeffrey Kubik and Amit Solomon, 2000, Security Analysts' Career Concerns and Herding of Earnings Forecasts, *RAND Journal of Economics* 31, 121-144.

Hong, Harrison, 2002, Analysts' Career Concerns and Biased Forecasts, Working Paper, Stanford University and Princeton University.

Irvine, Paul, Paul Simko and Siva Nathan, 2004, Asset Management and Affiliated Analysts' Forecasts, *Financial Analysts Journal* 60, 67-78.

Johnson, Timothy, 2004, Forecast Dispersion and the Cross Section of Expected Returns, *The Journal of Finance* 59, *1957-1978.*

Kim, Oliver, Steve Lim and Kenneth Shaw, 2001, The Inefficiency of Mean Analyst Forecast as a Summary Forecast of Earnings, *Journal of Accounting Research,* 39, 329-335.

Kim, Sok Tae, Ji-Chai Lin and Myron Slovin, 1997, Market Structure, Informed Trading, and Analysts' Recommendations, *Journal of Financial and Quantitative Analysis* 32, 507-524.

Krishnan, Ranjani and Donna Booker, 2002, Investors' Use of Analysts' Recommendations, *Behavioral Research In Accounting* 14, 129-156.

Kross, William, Byung Ro and Douglas Schroeder, 1990, Earnings Expectations: The Analysts' Information Advantage, *The Accounting Review*, 65, 461-76.

Lakonishok, Josef, Andrei Shleifer and Robert Vishny, 1994, Contrarian Investment, Extrapolation and Risk, *The Journal of Finance* 49, 1541-1578.

Loffler, Gunter, 1998, Biases in Analyst Forecasts: Cognitive, Strategic or Second-Best?, *International Journal of Forecasting* 14, 261–275.

McNichols, Maureen and Patricia O'Brien, 1997, Self-Selection and Analyst Coverage, *Journal of Accounting Research* 35, 167-199.

Michaely, Roni and Kent Womack, 1999, Conflict of Interest and the Credibility of Underwriter Analyst Recommendations, *The Review of Financial Studies* 12, 653-686.

O'Brien, Patricia and Ravi Bhushan, 1990, Analyst Following and Institutional Ownership, *Journal of Accounting Research* 28, 55-76.

Rao, Hayagreeva, Henrick Greve and Gerald Davis, 2001, Fool's gold: Social Proof in the Initiation and Abandonment of Coverage by Wall Street Analysts, *Administrative Science Quarterly* 46, 502-526.

Rock, Steve, Stanley Sedo and Michael Willenborg, 2001, Analyst Following and Count-Data Econometrics, *Journal of Accounting and Economics* 30, 351-373.

Stickel, Scott, 1991, Common Stock Returns Surrounding Earnings Forecast Revisions: More Puzzling Evidence, *The Accounting Review* 66, 402-416.

Stickel, Scott, 1992, Reputation and Performance Among Securities Analysts, *The Journal of Finance* 47, 1811-1836.

Stickel, Scott, 1995, The Anatomy of the Performance of Buy and Sell Recommendations, *Financial Analysts Journal* 51, 25-39.