Title
Three Essays in Empirical Finance

Author
Ahluwalia, Saurabh

Publication Date
2012

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Three Essays in Empirical Finance

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Management

by

Saurabh Ahluwalia

2012

ABSTRACT OF THE DISSERTATION

Three Essays in Empirical Finance

by

Saurabh Ahluwalia

Doctor of Philosophy in Management

University of California, Los Angeles, 2012

Professor Antonio Bernardo, Co-Chair

Professor Avanidhar Subrahmanyam, Co-Chair

How does information, dispersed among diverse geographic markets, customer segments and employees, get incorporated into the stock price? I have endeavored to create novel datasets that incorporate information from thousands of disaggregated investors, employees and customers. I utilize these datasets to construct aggregate measures of information and study how the information measures affect future returns and corporate events.

Specifically in the first chapter titled "Information Aggregation and Asset Prices", I utilize a unique data set based on Google Trends to construct a search index and use it to proxy for the information seeking behavior of retail investors. I find that abnormal search index predicts future buying pressure on the stock of a company. The portfolio with the highest increase in the search index has positive and significant alphas. The search index also predicts earnings surprises and is associated with the pre-earnings announcement drift. My results are robust to alternative specifications of CAR

windows, past returns, news coverage, information available to investors prior to the release of earnings numbers, and the information environment surrounding the earnings announcements. Overall, my results are in line with the hypothesis that retail investors' trades have information content relevant to stock prices.

In the second chapter titled "Effect of Employee Satisfaction on Earnings Surprises", I use a unique data set drawn from self administered employee surveys for 1495 US public corporations. I construct an Employee Satisfaction Index (ESI) and use it as a proxy for employee satisfaction. I find that ESI is higher for larger firms, high market to book ratio firms and firms that have low leverage. I also look at the effect of the changes in ESI on quarterly earnings announcements. I find that the changes in ESI are positively and (weakly) significantly related to the future quarterly earnings surprises. Moreover, the effect is stronger for companies that have higher information asymmetries and are more human capital dependent. The results are consistent with the theories that state that employees are insiders in a company and have information relevant to the future corporate performance. Moreover, consistent with human-capital centric theories I find evidence that the change in employee satisfaction has a greater effect on the performance of human-capital dependent companies.

In the final chapter titled "Private Equity Ownership and the Performance of Reverse Leveraged Buyouts", I study the effect of private equity exit on the target firm performance. Using a hand collected sample of 133 reverse leveraged buyout firms from 1997-2002, I examine the financial performance of the firms immediately before the IPO and up to four years after the IPO. I find that for three years after the IPO they continue to outperform their industries. However, performance deteriorates after the IPO. Cross-sectional regression at time of the IPO suggests that long term performance after the IPO is related to changes in ownership by the private equity sponsors and is

not related to changes in ownership by other insiders (all officers or directors who are not PE sponsors) or change in leverage. Even after the IPO, I find the positive relation between PE sponsor ownership and future performance continues. To establish causality between PE sponsor ownership and future performance I use the 2SLS-IV approach. The identifying instrument is the number of years since LBO and is a proxy for impatience of PE sponsors to free up their capital. I find that an IV regression finds a weakly significant relation between PE sponsor ownership and future performance.

The dissertation of Saurabh Ahluwalia is approved.

Jennie E. Brand

Mark J. Garmaise

Geoff Tate

Antonio Bernardo, Committee Co-Chair

Avanidhar Subrahmanyam, Committee Co-Chair

University of California, Los Angeles

2012

To My Mother, Pushpa Ahluwalia

To My Grandparents

TABLE OF CONTENTS

LIST OF TABLES

ACKNOWLEDGEMENTS

First of all, I would like to thank my co-chair Tony for his unfailing support and guidance. I feel very fortunate that I got an opportunity to work with Tony; a brilliant researcher, a caring teacher and most importantly a great man. I was also fortunate to have Subra as my co-chair. I gained immeasurably from his deep knowledge of the subject. I am extremely grateful to Geoff and Mark for numerous insightful discussions and continued guidance. I thank Jennie for her valuable critiques.

I thank Al Sheen for his invaluable comments and advise. I am grateful to numerous participants at the brown-bag seminars at UCLA for helpful comments. I am grateful to all the professors and students at UCLA for the countless ways they helped me. I am grateful to the staff for shielding me from paperwork and making administrative tasks a breeze. I am also grateful to staff at the Rosenfeld library for helping me with various database issues.

I thank Robert Hohman at Glassdoor for sharing with me the Employee data. I am grateful to Tarun Chordia for providing me with the Daily Buys and Sells data. I thank Jerry Cao and Josh Lerner for sharing their list of RLBOs with me.

Finally, and most importantly, I thank my mother and all my family for being there for me.

VITA

Education:

University of California at Los Angeles
John E. Anderson School of Management, PhD. Finance, 2012

Purdue University
Krannert Graduate School of Management, MBA (Finance and Marketing), 2002

Punjab University, Chandigarh, India
Punjab Engineering College, Bachelor in Engineering (Electronics), 1997

Work Experience:

Cendant, NJ, USA
Project Manager, Manager, Financial Planning and Analysis, 2002-2005

DRCI, MI, USA
Software Engineer, 1998-2000

Working Papers:

Information Aggregation and Asset Prices

Private Equity Ownership and Performance of Reverse Leveraged Buyouts

Work in Progress:

Effect of Employee Culture on Mergers and Acquisitions
with Albert Sheen (Harvard Business School)

Human Capital and Private Equity
with Albert Sheen (Harvard Business School)

Chapter I

Information Aggregation and Asset Prices

Abstract

Various theoretical models assume that information seeking behavior of investors has an impact on prices. However, it is very difficult to empirically test this, since the actual information acquisition process of the investors is unobservable. Using a unique data set from Google Trends, I construct a search index and use it to proxy for the information seeking behavior of retail investors. I find that abnormal search index predicts future buying pressure on the stock of a company. The portfolio with the highest increase in the search index has positive and significant alphas. The search index also predicts earnings surprises and is associated with the pre-earnings announcement drift. My results are robust to alternative specifications of CAR windows, past returns, news coverage, information available to investors prior to the release of earnings numbers, and the information environment surrounding the earnings announcements. Overall, my results are in line with the hypothesis that retail investors' trades have information content relevant to stock prices.

1.1 Introduction

Stock markets act as aggregators of information. Multitudes of investors seek information about different stocks that are traded in the market. It is the information seeking behavior of the investors that makes market prices efficient (Grossman (1976)). What is the mechanism by which the information flows into the market? How does information, dispersed among diverse geographic markets and customer segments, get incorporated into the stock price? Using a dataset based on Google searches, I try to answer these questions.

The process by which information is impounded in stock prices has been studied in detail by Grossman and Stiglitz (1980), Holthausen and Verrecchia (1990), Kim and Verrecchia (1997). An underlying prediction common to these models is that an investor will choose to become informed through information acquisition if the expected benefit of obtaining the information is greater than the cost of acquiring information. However, there has been little empirical research about the actual process used by investors to gather information and trade, since it is hard to observe an investor's information acquisition process. Using search data from Google Trends helps to circumvent this limitation.

I look at the Google searches of company stocks to proxy for the research being done on a company by a large community of investors[1]. Google Trends allows users to download search data for the keyword searches done through the Google search engine. Data is in the form of a search volume index that is based on the actual searches done for a particular keyword. I collect data for the S&P 1500 companies over five years from Jan 2004 to Dec 2008. I construct a weekly search

[1] According to the website hitwise.com, Google searches accounted for over 71 percent of all search queries performed in the US in 2010. Hence, Google searches for a topic can be considered representative of overall searches being done for a topic.

index that represents the volume of searches done for a stock. A high value of search index indicates that a large group of people are searching for information about a company[2]. The search can be motivated by a number of different reasons, such as employment or sales related inquiries. Such queries add noise as they are not investment related. To increase the power of my tests, I restrict attention to searches being done for the ticker symbol of the sample companies. Searches based on the ticker symbol come closest to ascertaining the motive that an individual may have for researching a company.

It is reasonable to assume that primarily individual investors are driving the search volume index since institutional investors will predominantly utilize sophisticated databases such as Bloomberg, and will tend to use outside sources less frequently. Second, searches by institutional investors will be a minuscule percentage of the overall searches done by all investors. Since Google Trends index does not include searches that are below a minimum search volume threshold, the search volume for keywords searched exclusively by institutional investors will most likely be set to zero. Finally, Da, Engelberg and Gao (2011) also find that there is a strong and direct correlation between changes in Google searches and trading by retail investors.

If information is indeed being aggregated by a multitude of investors, one would expect that the investors are trading based on the information. Hence, we would expect a link between the propensity to search and the subsequent trading volume of a stock. In the presence of short selling constraints for an individual investor, an increase in the propensity to search should be linked with the buy volume of the stock. To test this assumption I look at the buyer initiated orders for the

[2] There are many different ways investors can research a stock, such as reading analyst reports, reading newspapers, etc. However, in the past decade the internet has become the primary source of news and information in US. In a recent survey conducted in Jan 2011 by the Pew Project for Excellence in Journalism, (http://www.journalism.org) 34% of respondents said they read news online within the past 24 hours (in contrast to 31% who favored newspapers), and a full 41% said they get most of their news online (10% more than those who said they got most of their news from a newspaper).

stocks and find that an increase in searches has a positive and significant impact on future buy orders for shares of the company.

I find that change in the search index has a subsequent impact on weekly and monthly stock returns. Specifically, a portfolio that consists of stocks that experienced the highest increase in month over month search index earns an annualized risk adjusted excess return of 5.9%. To determine if the impact is permanent or temporary, I look at the returns for up to six months in the future and do not document any significant reversals. Results are similar when weekly changes in search index are used to form portfolios.

To further establish that it is the information aggregation that is driving the results, I look at events where stock specific information is disclosed at some future date. Even if each individual investor has imprecise information about the upcoming event, when the information is aggregated through the trades of many individuals, the resulting signal can still be very precise[3]. By looking at the interaction of the search index and cumulative abnormal returns around the event date, I can check the hypothesis about the information content of searches

In this paper, I use quarterly earnings announcements as a proxy for an information event. I find that investors tend to seek information prior to quarterly earnings announcements. An increase in the propensity to search for information about a company is linked to the future trading volume and predicts subsequent quarterly earnings surprises. Specifically, an abnormal change in the search index predicts three-day and five-day CAR around the earnings announcement. These results are consistent with Kaniel, Liu, Saar, and Titman (2008) and Griffin, Shu, and Topaloglu (2010), who

[3]Diamond and Verrecchia (1981) provide a theoretical framework under which in equilibrium, the private information of each trader contributes to the "noisy" aggregator of total information observed by all traders. As an illustration, of their model consider Schwert (1981), who finds a significant effect on the Standard and Poors 500 index of stock prices on dates when the consumer price index is announced. Each of the component prices of consumer price index is perfectly known by at least one investor. Hence, the information is known in the aggregate but not by any one individual, before the date of the CPI announcement.

find evidence that retail trading predicts returns around certain corporate events such as earnings and takeover announcements.

Since information is flowing constantly to the markets, we should expect that an abnormal change in search index will also have an effect prior to earnings announcements. I look at the CAR in the 30-day and 45-day sub-periods of pre-earnings announcement timeframes and find that the changes in the search index are positively related to pre-earnings announcement drift. I also look at post-earnings announcement CAR and find no evidence of reversals.

Further supporting the informational hypothesis, I find that the impact of searches is stronger for companies where information is scarce or hard to obtain. I find that the results are stronger for smaller companies in terms of both significance and magnitude of coefficients. Smaller companies tend to have greater information asymmetries and less news coverage. Moreover, they are tracked by fewer analysts and are not a major part of most institutional investors' holdings. Hence, individual investors who seek to research less widely followed companies have a better chance of earning fair rent for their efforts. My results are robust to alternative specifications of CAR windows, past returns, news coverage, information available to investors prior to the release of earnings numbers, and the information environment surrounding the earnings announcements.

I also consider three alternative hypotheses that may explain the results. First, it is possible that the abnormal search index is merely driven by news stories and or by efforts of a company through advertising expenses etc. However, the results are robust to the inclusion of news stories and measures of advertising expenses. Second, the "liquidity provision" hypothesis states that retail investors provide liquidity to institutional investors by taking the other side of institutional trades. Retail investors demand a premium for providing liquidity to institutional investors and thus exert temporary pressure on prices (Kaniel, Liu, Saar, and Titman (2008)). In such a case, we should see price reversals in matter of days. Third, it is possible that current period retail buying predicts more

retail buying in subsequent periods (auto correlated flow hypothesis). The buying pressure on stock leads to high future returns (Dorn, Huberman, and Sengmueller (2008) and Barber, Odean, and Zhu (2009)). Such prices increases are also subsequently reversed. However, I do not document any significant reversals for up to six-months.

My results are different from the paper by Da, Engelberg and Gao (2011) who use data from Google Trends for a sample of Russell 3000 stocks from 2004 to 2008. Their results provide support for Barber and Odean's (2008) naïve investors' buying pressure on the price hypothesis. I differ from them in terms data and construction of search index. Different from them, I look at effect of search index on buy sell volume and returns around earnings announcement. I discuss their paper in detail in the literature review.

This paper makes following primary contributions to the literature. I document the actual information seeking process of investors. I analyze the mechanism through which the information seeking behavior leads to trading and subsequent price impact in stock market. I find that retail investor trading serves as a conduit for information to flow from disaggregated sources to the market, thus making the market more efficient. Finally, I add to the growing amount of evidence that the retail investors can benefit by trading on private information (Kaniel, Saar, Titman (2008)).

This paper is organized as follows: section 2 briefly goes over the existing literature; section 3 describes the data and salient issues related with the data used in the paper; section 4 presents the results and section 5 summarizes and concludes.

1.2 Literature Review

This paper is closest to the paper by Da, Engelberg and Gao (2011) who use data from Google Trends for a sample of Russell 3000 stocks from 2004 to 2008. They find that the Google search index is correlated, yet, different from other proxies of investor attention such as market

capitalization, turnover, analyst following, and media attention. They also find that the Google search index predicts future retail volume and an increase in the search index contributes to large first-day returns and subsequent long run under-performance of IPOs. Their results provide support for Barber and Odean's (2008) naïve investors' buying pressure on the price hypothesis. They find some evidence that the search index predicts higher stock prices over the next two weeks. However, the price increase is subsequently reversed.

The current paper differs from Da et al. (2011) in the several important ways. First, while the focus of Da et al. (2011) is investor attention, the focus in this paper is the propensity of investors to seek information. Hence, the hypothesis, tests and conclusions in this paper are different from the ones in the Da et al. (2011) paper. Unlike Da et al. (2011), I look at the impact of the changes in the search index on information events such as earnings announcements. Moreover, I do not document any reversals in stock prices as is done by Da et al. (2011). Possible reasons for this are the difference in the sample size and the construction of the search index. While Da et al. (2011) collect their data with respect to an arbitrarily chosen Microsoft company ticker, I collect raw data for each ticker and then divide each weekly observation by an average of most common keywords related to the stock market. The advantage of the Da et al. (2011) method is that it allows cross-sectional comparisons as the search index of each company is calculated using MSFT ticker's search volume as the base. However, since the base is arbitrarily chosen their results are subject to the "Microsoft" and tech industry bias. I avoid the above-mentioned arbitrariness by working with the raw data (as was done by Choi and Varian (2009)).

Current paper is also close to Kaniel, Saar and Titman (2008) who look at short horizon returns subsequent to net buying by individual investors for 1,920 NYSE stocks from 2000 through 2003. They find stocks that are heavily bought by individuals in one week consistently outperform the market the following week. Kaniel, Saar and Titman (2008) propose that risk-averse individual

investors provide liquidity to institutions that demand immediacy. Hence, they suggest that as institutions sell to individuals prices fall one week and rebound the next. While I also look at the impact of retail investors on stock prices, I differ from KST in a number of key respects. I use a more primitive measure of retail investors' propensity to search for information as represented by the changes in the Google search index, to determine the impact of retail investors on the stock prices. I also differ from KST in terms of data and time frame under consideration. Additionally, I also test the impact of the changes in the search index by looking at CAR around the earnings announcements. Dorn, Huberman, and Sengmueller (2008) and Barber, Odean, and Zhu (2009) also find that retail buying predicts short horizon returns.

In a concurrent paper, Drake, M., D. Roulstone, and J. Thornock (2011 working paper) also find that when investors search for more information in the days just prior to the earnings announcement, the information content of earnings announcement is partially mitigated.

Google Trends data has been used in various disciplines and in a variety of settings. Ginsberg et al. (2009) show that one can estimate the level and outbreak of flu activity in different geographical areas in the US by looking at the search terms related to flu. They use the estimate to understand how the flu spreads and how can it be tracked. Choi and Varian (2009) use the Google search index as a proxy for the populations' demand for information. Choi and Varian (2009) find that the inclusion of the Google search index improves the model for predicting the monthly retail sales of automobiles and homes.

1.3 Data

1.3.1 Google Trends Data

Google Trends allows users to download data about volume of keyword searches done through the Google search engine. Data is available from Jan 2004 onwards and was first made

available to the public in Jan 2008. Google Trends data does not report the raw number of searches that were done for a particular keyword, but does report a search volume index, calculated via the following two step procedure: First, the query share is calculated by dividing the total query volume for a search keyword by the volume of the total number of queries for all the search keywords, and, second, the query share index is normalized so that it starts at 0 in January 1, 2004. Hence, the data at later dates is the percentage deviation from the query share on January 2004.

Search index data is available at the state level for the United States as Google is able to determine the geographical area of the user by decoding the IP address of the user making the query. Hence, using Google Trends one can download the search index that represents the search volume for a particular keyword for a given region at a particular point in time. I download the weekly (US only) search index for S&P 1500 Composite Index firms from Jan 2004 to Dec 2008. S&P Composite 1500 firms account for almost 90% of the U.S. market capitalization. I restrict the sample to S&P 1500 list of companies to make the task of data collection from the Google Trends and Google News Archive manageable.

The ticker symbol of each of the S&P 1500 firms was chosen as the keyword for which the search index data was downloaded. There are several advantages to using the ticker symbol. An individual may search for a company for a variety of different reasons such as looking for a job, checking a company product or some other non-investment related purpose. Such searches are devoid of any informational content for the stock price. Ideally, I would like to consider only those searches that are related to an investment purpose. At present, it is impossible to do so, since Google does not provide any contextual information for keyword searches. Hence, as a proxy for searches done for investment-related reasons, I look at the ticker symbol searches for the companies.

There are a few ticker symbols that stand for common words such as "WIN," "CAT," or "TRY." Hence, I exclude 178 tickers, from my sample, that also stand for common words. The search index is biased towards larger companies because Google Trends data truncates the searches if they are below a certain threshold. Tickers that do not get searched often or have search frequency so low that it does not meet Google Trends' minimum search volume requirement have their search index set at zero. Out of the 388,321 firm-week observations 196,837 are zeroes. Observations with the search index equal to zero are concentrated in predominantly small stocks. It is only natural that the small stocks that are less well-known and have fewer shareholders and employees will tend to be searched less often. I exclude the stocks for which the search index is zero for more than 90% of the weeks over the five-year period. This results in the exclusion of 391 stocks.

Data provided by Google Trends has a seasonal and growth component. From 2004 to 2008, both the market share of the Google search engine and the number of people accessing the internet has increased. To account for such mechanical properties of the data, I divide the weekly search index for each stock by an index composed of the average search index of the three most common investment related keywords (Google finance, Yahoo finance, and DJIA) searched by users. To ensure that my results are robust to the choice of the above keywords, I also used an index composed of the 10 most common keywords, with similar results.

1.3.2 Quarterly Earnings Announcements

Publicly owned companies are required by law to file quarterly earnings reports with the SEC. Many companies announce the earnings announcement dates in advance, but several companies miss these dates and report late. To determine the actual date of earnings announcements I use Compustat. Compustat records the earnings date as the date when the earnings report appears in the Wall Street Journal or other newspapers. However, the coverage of the Wall Street Journal is biased

towards well known large stocks, and it is possible that earnings announcement dates of smaller stocks are not accurate. To be certain that I have the correct earnings announcement dates, I cross check the earnings dates for smaller stocks against the Factiva database. I check for news stories to confirm that the date in Compustat is the actual earnings announcement date. In case of conflict, I use the date provided by the news story in Factiva.

1.3.3 Cumulative Abnormal Return and SUE

CAR centered on the earnings announcement is used as a measure of new information content (surprise) in the earnings announcement. Standardized Unexpected Earnings(SUE) can also be used as a proxy for the surprise/information in the earnings announcement. However, CAR does not suffer from the measurement error that SUE suffers from since for SUE one has to estimate the market's unobservable expectations of earnings. Moreover, in the earnings announcement management discloses other information apart from earnings. Firms provide extensive disclosure through financial reports, footnotes, management discussion of the results and the competitive environment, forecasts and other forward-looking information. In addition, many firms engage in conference calls where top executives of the firm present the last quarter's results and answer questions from investors and analysts. It is difficult to capture the full information content of earning's surprise by looking at SUE alone. This is consistent with the findings of Francis, Schipper and Vincent (2002) who find that earnings announcements increasingly serve as a conduit for information different from merely the earnings numbers.

However, I use Standardized Unexpected Earnings (SUE) as a proxy for market expectations surrounding the earnings announcement. Analyst forecast data is from I/B/E/S in WRDS. SUE is defined as: (Actual EPS - Expected EPS) / (Standard deviation of analyst forecasts for that quarter) Actual EPS data is from Compustat. Expected EPS is the mean of the analyst forecasts prior to the

announcement. If an analyst has made multiple forecasts in this period, only the most-recent forecast is used. If there are less than two forecasts for that quarter, it leads to exclusion of the firm for that quarter. Absolute magnitude of SUE measures the degree of earnings surprise, while the sign of SUE signifies if the actual earnings were above or below the consensus forecast. That is, a positive SUE implies that the actual values of EPS came out to be above the analysts estimates and a negative SUE signifies that the earnings numbers came below the mean analyst estimate.

1.3.4 News Data

I collect weekly news data for S&P 1500 companies from 2004 to 2008. I utilize Google News Archive to search for news (in English only) related to the ticker and company name of a firm. For each week, from the beginning of 2004 to the end of 2008(260 weeks for each of the S&P 1500 companies), I search for news stories for the company ticker and the company name. The number of news stories for each firm, for a particular week, is recorded and constitutes the variable "News Volume".

There are various advantages to using Google News data. First, it is probably the most comprehensive news data available. It is better than DJI and Lexis Nexis news data source since it includes many news sources that other databases do not include. For example, many online news sources and local newspapers are not part of DJI and Lexis Nexis. Prominent journal papers that look at the news data typically only look at the top three to five newspapers. Moreover, Google News data is most relevant to this paper as it captures the information available to an online investor.

1.3.5 Volume Data

I also look at the impact of the searches for a company ticker on the subsequent trading volume of the company's stock. I use two primary measures of volume: daily buys and sells volume, and daily buys and sells dollar volume. Daily buys and sells volume data was kindly provided by Tarun Chordia. Daily buys and sells are calculated using data from the NYSE's TAQ database. The trades for each stock are signed based on the Lee and Ready (1991) algorithm. The quote rule and tick rule are used to identify the trades as buyer-initiated or seller-initiated. If the trade price is above the midpoint of the most-recent bid-ask quote, the trade is classified as buyer initiated. Additionally, if the trade price is above the last executed trade price, the tick rule classifies the trade as buyer-initiated.

1.4 Empirical Results

1.4.1 Summary Statistics

Table I-1 shows the distribution of the S&P 1500 companies by industry for the year (2006). Manufacturing has the most representation in the sample while Food Products has the least representation. Table I-2 presents salient statistics for the sample firms. Table I-3a shows the summary statistics for the search index, averaged monthly, for each year. There are clear outliers in the data (exceeding the mean by over 20 standard deviations) and the results in this table and the subsequent tables are winsorized at 1% and 99% level. Also shown is the average of the month over month differences. The mean change is slightly positive for each year and varies between -3.2 to 3.5.

1.4.2 Weekly Portfolio Rebalancing

In the following section, I analyze the performance of the portfolios formed based on the changes in the search index. I want to test 1) If the portfolio with greater positive change in the prior search index earns higher excess returns. 2) If the performance of the portfolio can be

explained by the style tilts or factor loadings. 3) If there are any reversals in the future time periods. Since my sample period is small (five years), I look at the weekly returns of the portfolios formed on bases of change in the search index. I define change in the search index as the difference in the search index for week 5 minus the average search index for week 1 to week 4. Each week, I divide the stocks into five portfolios based on the changes in the search index. Hence, portfolio 5 at end of week *w* will have stocks that had experienced the most increase in the search index from the prior four week average search index Similarly, portfolio 1 at end of week w will have stocks that had experienced the most decrease in the search index from the prior four week average search index Every week the portfolios are rebalanced and stocks are assigned to new quintiles. Excess returns are calculated for week w+1 by subtracting the risk-free rate. Excess returns are regressed against the risk factors, as specified by equation (1) below:

$$R^e_{i,t} = \alpha_{i,t} + \beta_i R^e_{m,t} + s_i SML_t + h_i HML_t + \varepsilon_{i,t} \qquad (1)$$

$$i = 1,2,...,5$$
$$t = 1,2,...,T$$

Here $R^e_{i,t}$ is the value-weighted return of the portfolio i in week t in excess of the benchmark risk free rate. α is the primary variable of interest and it shows if the trading strategy is generating risk adjusted excess returns. $R^e_{m,t}$, SML_t, HML_t are respectively the excess return on the market, size and value factors taken from Ken French's website. To account for serial correlation, I use the Newey West correction for standard errors.

Table I-4 shows the results of regressing the weekly excess returns on the Fama-French (1993) risk factors. The results show that if one were to buy the stocks in the highest quintile one week after the increase in the index, one can earn a weekly alpha of 0.132%, significant at 1% level. The 132 basis point weekly spread is economically large and translates into an annualized return of

6.86%. It should be noted that the trading strategy of buying the stocks in portfolio 5 and shorting the stocks in portfolio 1 does not yield a significant alpha (column (6)). It is difficult to give a meaningful interpretation to the decrease in the search index, and it is not clear if the decrease in the search index should have any effect on the stock returns.

The above results show that an increase in the search index is associated with a subsequent positive price impact on stocks. The price increase could be caused by various reasons. One possibility, that is consistent with the "information hypothesis", is that the increase in the search index leads to information flow into the stock prices. An increase in searches for a firm indicates that retail investors are indulging in information-seeking behavior and are actively seeking information about the company. This behavior could have been triggered by some private information or observation about the firm. If their research sheds a positive light on the company, they end up buying it. However, if the investors uncover something negative, they tend not to do anything since retail investors tend not to short-sell.

It is also possible that, consistent with the "liquidity provision" hypothesis, the price run-up occurred simply because the retail investors were providing short-term liquidity to the institutional investors. As argued by KST (2008), one would expect price reversals in a matter of days once the short-term liquidity needs of the institutional investors are satisfied. Table I-4 shows the results for holding the portfolio at future weeks up to two months after the portfolio formation date. I do not document any reversals expected to take place in the future. An absence of reversals indicates that the price impact is not short-lived, as will be predicted by the "liquidity provision" hypothesis. If retail investors were simply providing short-term liquidity to the institutional investors, one would expect price reversals in a matter of days, as shown in KST (2008).

1.4.3 Monthly Portfolio Rebalancing

As a further robustness check, I rerun the above analysis on a monthly basis. I divide the stocks into five portfolios based on changes in the search index from one month to the next. Hence, portfolio 5 at the end of month m will have stocks that have experienced the most increase in the search index from month m-1 to month m. Similarly, portfolio 1 at the end of month m will have stocks that have experienced the most decrease in the search index from month m-1 to month m. Every month the portfolios are rebalanced and stocks are assigned to new quintiles based on the change in the index from previous month. Excess returns are calculated for month m+1 by subtracting the risk free rate. Excess returns are regressed against the risk factors, as specified by equation (1) below:

$$R^e_{i,t} = \alpha_{i,t} + \beta_i R^e_{m,t} + s_i SML_t + h_i HML_t + m_i MOM_t + \varepsilon_{i,t} \qquad (2)$$

$$i = 1,2,...,5$$
$$t = 1,2,...,T$$

Here $R^e_{i,t}$ is the return of the portfolio i in month t in excess of benchmark risk free rate.

α is the primary variable of interest and it shows if the trading strategy is generating risk adjusted excess returns.

$R^e_{m,t}$, SML_t, HML_t, MOM_t are respectively the excess return on the market, size, value and momentum factors taken from Ken French's website. To account for serial correlation, I use the Newey West correction for standard errors.

Table I-5 shows the results of the monthly regressions. Alpha for portfolio 5, the highest positive search index change portfolio, is significant at a 10% level. Based on one month returns, the alpha translates into an annualized excess return of about 5.9%. The weekly alpha is about one percent higher than the alpha for monthly returns. However, in case of weekly portfolio rebalancing, the effective returns are diminished by the transaction costs accruing because of the increased

trading frequency. I also look at the performance of the portfolio up to six months in advance and I do not document any significant price reversals.

It is natural to ask why the market is not able to discern the effect of the increase in the search index on the returns. One of the reasons is that the information on aggregated searches has been public only since Jan 2008 and the market had no way of accessing this information prior to this period.

To further check the information hypothesis, I look at the information dissemination through quarterly earnings announcements and its relation with the search index.

1.4.4 Quarterly Earnings Announcements

One of the ways to test whether the search index has any price impact and has information pertaining to stock prices is to look at the events where there is a dissemination of information relevant to stock prices, such as through a regulatory change, a court ruling, or by the company itself as in its earnings announcements. By studying the behavior of prices and the search index around the event one can test the above-mentioned hypothesis about the information flow in stock prices.

In this paper, I focus on the quarterly earnings announcements for a variety of reasons. All publicly traded companies are required by law to have four quarterly earnings announcements every year at periodic intervals. This gives me a potential 1500*5*4 or about 30,000 firm-quarter data points to test my hypothesis. Since all public companies have to make these announcements, sample selection biases are considerably reduced. In case of FDA approvals or court rulings, the sample will be much shorter and can be biased towards certain industries. Moreover, periodic analyst forecasts before the earnings announcements allow me to proxy for market expectations and uncertainty surrounding the announcement and help me to develop a richer set of tests.

Table I-6. shows the salient summary statistics and distribution for the primary variables used in the earnings related regressions. SVI is the average of the one month search index prior to the earnings announcement. Five calendar days prior to the earnings announcement are excluded so that there is no effect of earnings leakages on the search index. NewsVolume is the total number of news stories that include the name of the company and its ticker symbol, in the one month period prior to the earnings announcement.

News coverage is heavily skewed towards larger companies. While the top one percentile of companies have an average of 84 or more news stories in the one month time period prior to the earnings announcement date, almost half the sample firms have on average less than two news stories in the same time period. Mean and median three-day and five-day CAR centered on the earnings announcement date is slightly positive. One would expect the CAR to be zero since market should not systematically under- react to the earnings information. However the variance of all the three specifications of CAR is quite high, hence the positive mean and median are not significant in any of the specifications.

For most of the companies, the earnings announcement process is an opportunity to aggregate sales and profit numbers and share the results with the public. Earnings surprise indicates that the company performed better or worse than the market expectations. Earnings surprise is the new information that while present in disaggregate form among the diverse employees and customers of the company, gets revealed to the market only when the information has been aggregated by the management of the company and shared with the public on the day of the earnings announcement.

Changes in the search index also mirror the above aggregation of information. A multitude of customers and employees of the company observe the company's products and services first hand. When they get excited about the company, they may decide to research the company further before buying its stock.

Hence, we can expect that the search index of a firm will be associated with future earnings surprises of the firm. As a proxy for the earnings surprise, I utilize the cumulative abnormal return (CAR) centered on the earnings announcement. CAR centered on earnings announcement reflects the impact of new information that is disclosed by the earnings announcement after taking into account market expectations. Specifically, I use the three-day and five-day abnormal returns around the earnings announcements. General specification for the regression equation is -

$$\sum_{x}^{y} CAR_{i,t} = \alpha_{i,t} + FirmDummies + TimeDummies + \beta_1 ChangeSearchIndex_{i,t-1} + ControlVariables_{t-1} + \varepsilon_{i,t} \quad (3)$$

Number of firms i =1, 2,… N

Number of quarterly earnings announcements t =1, 2,…Nq

$\sum_{x}^{y} CAR_{i,t}$ is the cumulative abnormal return for firm i from day t+x to *day* t+y and is also represented by the suffix [x,y]. Abnormal return is defined as the return of the stock less the value weighted market return from CRSP. The difference is summed up over each of the trading days in the period.

The search index for a firm i is the average search index from time period t+a to t+b. I exclude the five days prior to the earnings announcement to mitigate the effects of insider trading or leakage of earnings information to the market. ChangeSearchIndex is the difference between average search index from day -35 to -5 and the average search index from day -66 to -36, where day 0 is the day of earnings announcement. Hence, I am testing if for a given firm, the change in the level of search index can predict future abnormal returns.

Table I-7 shows the results for the above specification. Specification (1) in Table I-7 shows that the one-month search index prior to the earnings announcements predicts five-day CAR at a

5% level of significance. A greater increase in the search index is associated with a larger earnings surprise. It is also possible that the results are being driven by investors who get overly excited when they see a stock price going up prior to the earnings announcement and then trade excessively, pushing the price up temporarily. Thus, it is the extreme past returns (Barber and Odean, 2008) that are driving the results. I control for the possibility that past returns may be driving the results by including a lagged two-month return. Results are robust to inclusion of one to six months of lagged returns. Specifications in Table I-7 also include other control variables commonly used in the earnings announcement regressions (Kothari 2001).

It is possible that the search index is capturing the information content in the analyst reports. Financial analysts collect information from diverse sources, study past and current performance of the companies they follow, and make forecasts about the company. The search index could be merely substituting for the information and news stories about the company that analysts use to predict earnings, and is already available to investors. It is important to check if the results hold, even while controlling for the analysts forecasts about the earnings. Analysts tend to follow companies closely and their reports should be able to capture part of the information contained in the search index. I find that in case of earnings surprises, the effect is mitigated by the addition of proxy for the analyst information prior to the earnings announcement (SUE (standardized unexpected earning)) indicating that analysts are indeed relaying to the markets part of the information that is contained in the search index.

Table I-7 (specification 2) shows the results once I include the SUE in the regression equation. To control for the information environment and uncertainty surrounding the earnings announcement, I also include the standard deviation of analysts forecast prior to the announcement. SUE is highly significant, as one would expect, and it does have an impact on the search index in terms of diminishing the magnitude of the coefficient of the search index variables. However, the

search index retains most of its explanatory power and is still significant at the 5% level. The above results indicate that the search index has some of the information content that is being captured by the analysts.

To check if prior news coverage is driving the results, I include the variable NewsVolume in specification 3. NewsVolume is the number of news stories that mention the firm in the one month period prior to the earnings announcement date. The coefficient of NewsVolume variable is highly significant, indicating that, for the same firm, the periods in which it gets higher news coverage are associated with higher CAR around the future earnings announcement date.

Market to book ratio (MTB) and lagged returns have negative and significant coefficients. It seems that in the periods where there is a run up in the stock price prior to the earnings announcements there is a negative effect on the CAR around the date of an earnings announcement. SG&A proxies for the advertising expense incurred by a firm. A higher number of advertisements may lead consumers to research a company. SG&A is included to control for the possibility that it is merely advertising that is driving the results.

Results are economically significant. A one-unit increase in the search index is associated with a 80 basis points increase in the five-day CAR. To control for any spurious effects that can be contaminating the stock price in the five-day window, I also look at a shorter window specification. Specifications 4, 5 and 6 detail the results for the three-day CAR, centered on earnings announcement. Results are robust to the alternative specifications of the event window size.

As a further test of the information hypothesis, I consider the possibility that investors get overly excited about the company for reasons unrelated to the information, which causes the subsequent run-up in prices prior to and around the earnings announcement. In such a scenario, one can expect reversals to occur in the stock price after the announcement. I test for price reversals up to 45 days after the announcement date. Table I-7 (specifications 7 and 8) show the result of the

regression of the 40-day CAR from day +5 to day +45 and the 30-day CAR from day +5 to day +35 on the one month prior search index, respectively. The five calendar days subsequent to the earnings announcements are excluded in order to account for any abnormal price moves immediately after the announcement.

The coefficient of the search index prior to the earnings announcement is positive for the 30-day CAR and negative for the 40-day CAR. However, the coefficient is insignificant in both the specifications and it does not appear that any significant price reversals are taking place.

Models by Grossman and Stiglitz (1980), Holthausen and Verrecchia (1990), Kim and Verrecchia (1997) predict that an investor will choose to become informed through information acquisition if the expected benefit of acquiring the information is greater than the cost of acquiring information. Hence, if indeed it is the information content of searches that is driving the results, then we can expect that the results will be stronger for the firms where the information is less readily available. In general, smaller firms have a lower analyst following and the smaller firms tend to be covered less frequently by journalists. We should expect that the impact of the searches will be stronger for the companies where information is scarce or hard to obtain.

To test if the impact of searches is higher for smaller firms, I rerun the regression in Table I-7 separately for small and large firms. The sample is divided into three bins based on size. "Small Firms" constitute the firms in the smallest size bin while the "Large Firms" constitute the firms in the largest size bin. Specifications 1, 2 in Table I-8 show the results of regression of the five-day cumulative abnormal return on the lagged Search Index. Abnormal return is calculated by subtracting the value weighted market return from the stock return. SVIn is the average one month search index from day -35 to -5, where day 0 is the day of the earnings announcement. SVIo is the average one month search index from day -36 to -66. NewsVolume is the number of news stories that mention the firm in the (-36 to -66) time period. For specifications 5 and 6 CAR is calculated

from day (+5 to +30) and for specifications 7 and 8 CAR is calculated from day day (+5 to +45), where day 0 is the day of the earnings announcement.

I find that the results are stronger in terms of both the significance and magnitude of coefficients for smaller companies. Smaller companies tend to have greater information asymmetries, have less news coverage, are followed by fewer analysts, and are not the major portion of most institutional investors' holdings. Hence, individual investors who seek to research less widely followed companies have a better chance of earning a fair rent for their research efforts.

The above results support the information hypothesis that the increase in the search index has information content that has a permanent effect on the price and can predict earnings surprises. For example, an earnings surprise can indicate that, in the previous quarter, the company did something unusual, such as, introducing a product or improving its customer service that resulted in profits beyond the expectation of the market. It is likely that customers and employees experience this change firsthand, and this prompts them to seek information about the company. The increase in the search index proxies for the information seeking behavior and helps to aggregate information from diverse sources, much like what the management of the company does at the end of the quarter when it compiles its earnings numbers from different stores or plants in various markets.

As a more stringent test of the information aggregation story, I also look at the pre-earnings drift to determine if, in fact, the information has been gradually flowing to the market prior to the earnings announcement.

1.4.5 Pre- Earnings Announcement Drift

The pre-earnings announcement period allows me to test the hypothesis that the search index is related to the gradual flow of information to the stock market prior to the earnings

announcement date. Pre-earnings announcement drift is defined as the CAR n days prior to the earnings announcement date. If a higher number of searches led to net buying behavior, we should expect that the more a stock is searched before its earnings announcement the higher will be its expected pre earnings announcement drift.

I look at the 30 calendar day CAR from day -35 to day -5 and the 40 calendar day CAR from day -45 to day -5, where 0 is the earnings announcement date. I exclude the last five days prior to the earnings announcement to mitigate the effects of insider trading or leakage of earnings information to the market. Abnormal return is defined as the return of the stock less the value weighted market return from CRSP. The difference is summed up over each of the trading days in the 30 and 40 days prior to the announcement.

The independent variable of interest is the prior month search index (from day -36 to day -66). Specification 1 in Table I-9 reports the result of the panel regression of CAR on the search index with firm and time fixed effects. A higher search index predicts a higher pre-earnings announcement drift. Since I am using firm fixed effects, the independent variable is the change in the search index from its long term mean. Hence, I am testing if, for a given firm, the change in level of the search index can predict future abnormal returns. A one-unit change in the search index is associated with an increase of 0.83% increase in the 30-day CAR prior to the earnings announcement. Specifications 2 and 3 also control for common independent variables as well as lagged returns and news coverage.

It is possible that there is a third variable that is related to both searches and future returns, that is causing the results. As Barber Odean (2008) show that high past returns can lead to increased investor attention, and this may cause the searches to rise in the future. Future returns can also be higher because of momentum generated by past returns. To control for such a scenario, I include in the regression the past two months' returns. The results are robust to inclusion of past returns up to

the previous six months. It does not appear that the search index is being driven by past momentum in the returns.

It is possible that a stock gets lots of news coverage and investors are merely buying or selling the stock based on the coverage alone. I include the prior month news coverage variable "NewsVolume" to control for such a scenario. NewsVolume is the number of news stories that mention both the stock ticker and the company name in the prior one month time period. Results are robust to the inclusion of the above variables.

1.4.6 Trading Volume

In the above analysis, I contend that market prices are informative since investors are trading based on their information. In such a case, we should see the effect of the search index on the trading volume[4]. Hence, to link the search index with actual trading activity, I look at the effect of the search index on future trading volume. It can be argued that unsigned volume merely represents the change in ownership of shares; that is to say for every buyer, there is a seller. However, on average, stock prices tend to rise in periods of high volume and fall in periods of low volume, as shown by Karpoff (1987). If an increase in the search index leads to an increase in buying activity prior to the earnings announcement, we should also see a subsequent increase in trading volume. Specifically, given the short selling constraints faced by individual investors (Barber Odean 2008) the effect should be stronger for buy orders. I use three proxies for the buying pressure on the stock to determine if a higher level of the search index is associated with more intense buying activity at a later date.

[4] Even when the information observed by investors is identical, there can be an effect on volume because of differences in opinions of different class of investors about the relation between the news and asset value, Harris and Raviv (1993).

In Table I-10, the daily buys and sells volume data was used. Daily buys and sells are calculated using data from the NYSE's TAQ database. The trades for each stock are signed based on the Lee and Ready (1991) algorithm. Daily buys and sells were considered for the period (-5 to -45) and (-5 to -30) calendar days before the earnings announcement. The five days prior to the earnings announcement were excluded to account for any insider trading activity prior to the earnings announcement. Sample size is reduced since I have the buys and sells data for only a subsample of S&P 1500 firms that are listed in the NYSE.

The dependent variable in the specification (1) is defined as follows- (sum of daily buys from (-45 to -5) / {(sum of daily sells from (-45 to -5) + (sum of daily buys from (-45 to -5)}. For specification (2), the dependent variable is the (sum of dollar daily buys from (-45 to -5)/ {(sum of dollar daily sells from day (-45 to -5)} + (sum of dollar daily buys from (-45 to -5)}. It is possible that the results are being influenced by a heavy buying or selling day in the time period (-45 to -5). Since I am summing all the buy orders in the time period, the results may have been influenced by one or two intense buy volume days. In specifications (3) and (4), I use the average of the buys ratio for each day to mitigate the effects of intense buying or selling days.

For specification (3), the dependent variable is defined as the average of the daily buys ratio from day (-45 to -5). The daily buys ratio is defined as the ratio of daily buys divided by the sum of buys and sells for the day. The dependent variable for specification (4) is defined as the average of the dollar daily buys ratio from day (-45 to -5). The daily dollar buys ratio is defined as the ratio of dollar daily buys divided by the sum of dollar buys and dollar sells.

Table I-10 shows that the increase in the prior month search index score (from day -46 to -76) is positively related to the subsequent increase in the buy trading volume. In each of the specifications, coefficients are significant and economically large. For robustness, an alternate window from day (-5 to -30) was also considered and the results remain robust to the alternate choice of the window and

are shown in the specifications (5) to (8). These findings further support the hypothesis that the increases in search activity translates into the buying behavior of investors.

1.5 Conclusion

Using online searches for a ticker symbol on the internet as a proxy for investors' information seeking behavior, I look at the impact of the searches on stock prices. Data from Google trends allows me to proxy for the active interest of millions of investors in a stock. It is possible that some event or information may have triggered the interest which leads investors to further research a company. Since the internet is the biggest news and information source available to a retail investor, it is expected that one will seek stock information online.

I show that the search index-representative of volume of searches being done for a particular stock- has a permanent price impact. To test whether an increase in search activity predicts a future price increase that leads to incorporation of information in stock prices and subsequent price rise, I form five portfolios based on the month over month change in the average search index. I find that the portfolio with the largest increase in the search index has a significant risk adjusted alpha that translates into an annualized return of 5.9%. The price impact seems permanent and I do not document any reversals in the price over subsequent periods of up to six months. The results are similar when portfolios are rebalanced each week based on the week over week changes in the search index.

To further establish that the increase in search index is related to informational reasons, I look at the association between the search index and returns around information events such as quarterly earnings announcements. Using cumulative abnormal returns around the quarterly earnings announcements I find that the search index can predict pre-earnings announcement drift as well as

the earnings surprise of a firm. Results are stronger for the firms where the information is less readily available. In general, smaller firms have a lower analyst following and journalists cover the smaller firms less frequently. Consistent with the notion that investors should get higher returns if they uncover hard to obtain information for firms, I find that the impact of the increase in search index is stronger for smaller companies.

In the above analysis, I contend that market prices are being made informative since investors are acting on their information and are trading. In such a case, we should see the effect of an abnormal search index on the trading volume. Specifically, given the short selling constraints faced by individual investors (Barber Odean 2008) the effect should be stronger for buy orders. I use three proxies for the buying pressure on a stock to determine if a higher level of the search index is associated with more intense buying activity at a later date. Each of these measures is based on the actual buys and sells data from the TAQ database. I find evidence that the search index is associated with future buying pressure on the stock.

To summarize, I find support for the information hypothesis, that retail investor information seeking behavior as captured by search index has information content that has a permanent price impact.

Table I-1. Summary Statistics (A)

This table shows the industry distribution of S&P 1500 firms. The sample period is from January 2004 to December 2008.

SIC Codes	Industry Description	Number Of Firms	%
0-1000	Food Products	7	0.47%
1000-1999	Mining and Construction	78	5.20%
2000-2999	Consumer Products	216	14.40%
3000-3999	Manufacturing	375	25.00%
4000-4999	Utilities and Transportation	148	9.87%
5000-5999	Wholesale and Retail	172	11.47%
6000-6999	Financial Services	268	17.87%
7000-7999	Personal & Business Services	159	10.60%
8000-8999	Miscellaneous	77	5.13%
Total		1500	100.00%

Table I-2. Summary Statistics (B)

This table shows the summary statistics for the sample S&P 1500 firms for the year 2006. The sample period is from January 2004 to December 2008. Assets is total assets (Compustat item actq). MktValue is market value of common equity and is defined as common shares outstanding times the quarter end price (Compustat cshoq*prccq). LnSale is natural log of sales (Compustat item saleq). Cashflowat is the sum of income before extraordinary items and depreciation and amortization divided by lagged assets value (Compustat (ibq+dpq)/l.atq).

RetonAssets is income before extraordinary items divided by lagged assets value (Compustat ibq/l.atq). MTB is market to book ratio and is defined as the sum of total assets and market value of equity minus the book value of equity divided by total assets (Compustat (atq-seqq+cshoq*prccq)/atq). LevBook is defined as long-term debt plus current liabilities divided by long-term debt plus current liabilities plus book value of common equity (Compustat (dlcq + dlttq) / (dlcq + dlttq + ceqq)). LevMkt is defined as long-term debt plus current liabilities divided by long-term debt plus current liabilities plus market value of common equity (Compustat (dlcq + dlttq) / (dlcq + dlttq + cshoq*prccq)).

	Mean	Std. Dev.	Min	Max	NumObs
Assets	18,417	108,087	4.07	1,668,498	1500
MktValue	9,115	25,544	21.10	368,223	1494
LnSale	5.812	1.781	-3.863	11.286	1500
Cashflowat	0.025	0.030	-0.313	0.280	1348
RetonAssets	0.015	0.027	-0.314	0.280	1494
MTB	2.235	1.642	0.582	16.238	1494
LevBook	0.294	0.244	0.000	1.011	1383
LevMkt	0.169	0.178	0.000	0.895	1373

Table I-3. Summary Statistics (C)

This table shows the summary statistics of the monthly search index for the sample S&P 1500 firms. The monthly search index is the average of the weekly search index for the month. The sample period is from January 2004 to December 2008.

Monthly Search Index

Year	Obs.	Mean	Std. Dev.	Min	Max
2004	9946	0.580	0.513	0.000	3.063
2005	9923	0.710	0.499	0.000	3.575
2006	9897	0.848	0.463	0.000	3.480
2007	9884	1.010	0.492	0.000	3.538
2008	9802	1.155	0.582	0.000	3.575

Month Over Month Change in Search Index

Year	Obs.	Mean	Std. Dev.	Min	Max
2004	9110	0.008	0.180	-2.650	2.930
2005	9911	0.012	0.205	-2.650	2.650
2006	9880	0.010	0.226	-2.800	3.350
2007	9868	0.017	0.270	-3.200	3.525
2008	9760	0.003	0.214	-2.108	3.050

Table I-4. Weekly Risk-Adjusted Returns of the Portfolios

This table shows the results of weekly regressions of the excess portfolio returns on the Fama-French(1992) factors, Rm-Rf, HML and SMB. Excess returns are calculated by subtracting the risk free rate from the value-weighted portfolio returns. Rm-Rf is the value-weighted market return on all NYSE/Amex/Nasdaq firms minus the risk-free rate. SMB (small minus big) is the difference each week between the return on small- and large-capitalization firms. HML (high minus low) is the difference each week between the return of a portfolio of high book-to-market stocks and the return of a portfolio of low book-to-market stocks. Each week the stocks in the S&P 1500 index are assigned to one of the five portfolios based on the weekly change in the search index. Stocks experiencing the largest week over week increase are assigned to portfolio five, while the stocks that experience the largest week over week decrease are assigned to portfolio 1. Specifications 7 to 12 show results for portfolio 5 for subsequent weeks. For example, p5 W+2 constitutes one week returns, two weeks after the date of portfolio formation. Robust standard errors are in parentheses. The sample period is from January 2004 to December 2008. . ***, **, * denote 1%, 5% and 10% significance levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	potrfolio 1 (p1) - largest decrease in SVI	**portfolio 2**	**portfolio 3**	**portfolio 4**	**potrfolio 5 (p5) - largest increase in SVI**	**p5-p1**
Alpha	0.00035	0.00043	0.00032	0.00008	0.00132***	0.00097
	(0.0005)	(0.0005)	(0.0007)	(0.0006)	(0.0005)	(0.0006)
Rm-Rf	0.855***	0.878***	0.794***	0.868***	0.885***	0.030
	(0.0487)	(0.0314)	(0.0784)	(0.0345)	(0.0317)	(0.0371)
SMB	0.030	0.009	0.251	-0.081	-0.001	-0.030
	(0.0778)	(0.0542)	(0.1265)	(0.0628)	(0.0458)	(0.0705)
HML	-0.195***	-0.213***	-0.029	-0.163**	-0.139**	0.055
	(0.0490)	(0.0592)	(0.0789)	(0.0711)	(0.0589)	(0.0629)
N	292	292	292	292	292	292
R Sq.	0.91	0.91	0.83	0.85	0.89	0.02

	(7)	(8)	(9)	(10)	(11)	(12)
	p5 (w+2 week)	**p5 (w+3 week)**	**p5 (w+4 week)**	**p5 (w+8 week)**	**p5 (w+9 week)**	**p5 (w+10 week)**
Alpha	-0.00030	0.00074	0.00044	0.00020	0.00076	0.00038
	(0.0005)	(0.0005)	(0.0004)	(0.0005)	(0.0005)	(0.0005)
Rm-Rf	0.878***	0.872***	0.883***	0.813***	0.838***	0.926***
	(0.0442)	(0.0249)	(0.0278)	(0.0587)	(0.0540)	(0.0341)
SMB	0.038	-0.058	-0.064	0.059	0.053	-0.004
	(0.0760)	(0.0460)	(0.0433)	(0.0932)	(0.0889)	(0.0603)
HML	-0.162***	-0.155**	-0.114**	-0.071	-0.061	-0.156**
	(0.0487)	(0.0512)	(0.0525)	(0.0454)	(0.0483)	(0.0554)
N	292	292	292	289	288	287
R Sq.	0.91	0.89	0.91	0.88	0.89	0.91

Table I-5. Monthly Risk-Adjusted Returns of the Portfolios

This table shows the results of monthly regressions of portfolio returns on the FF three factors (Rm-Rf, SMB, HML) and Carhart's momentum factor. Excess returns are calculated by subtracting the risk free rate from the value-weighted portfolio returns. Rm-Rf is the value-weighted market return on all NYSE/Amex/Nasdaq firms minus the risk-free rate. SMB (small minus big) is the difference each month between the return on small- and large-capitalization firms. HML (high minus low) is the difference each month between the return of a portfolio of high book-to-market stocks and the return of a portfolio of low book-to-market stocks. Each month the stocks in the S&P 1500 index are assigned to one of the five portfolios based on the month over month change in the search index. P5 is the portfolio that includes stocks that have had the largest increase in the search index from the previous month. Similarly, P1 is the portfolio that includes stocks that have had the largest decrease in the search index from the previous month. Specifications 7 to 11 show results for portfolio 5 if held for additional months. For example, p5 (m+2) looks at monthly returns two months in future. Dependent variable is the value-weighted return less the risk free rate. Standard errors are in parentheses. The sample period is from January 2004 to December 2008. . ***, **, * denote 1%, 5% and 10% significance levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	potrfolio 1 (p1) - largest decrease in SVI	**portfolio 2**	**portfolio 3**	**portfolio 4**	**potrfolio 5 (p5) - largest increase in SVI**	**p5-p1**
Alpha	0.0029	-0.0007	0.0001	0.00022	0.0049**	0.002
	(0.0021)	(0.0017)	0.0023	(0.0016)	(0.0019)	(0.0030)
Rm-Rf	0.88***	0.83***	0.95***	0.82***	0.77***	-0.11
	(0.0640)	(0.0520)	0.06	(0.0420)	(0.0530)	(0.0800)
SMB	-0.12	-0.0033	-0.236*	-0.018	0.103	0.22
	(0.1340)	(0.1070)	0.124	(0.0980)	(0.0973)	(0.1800)
HML	-0.13	0.013	0.034	0.266***	-0.12	0.011
	(0.0850)	(0.0760)	0.114	(0.0790)	(0.0720)	(0.1180)
Mom	-0.059	-0.062	-0.179***	-0.119*	-0.093	-0.034
	(0.1000)	(0.0624)	0.06	(0.0576)	(0.0590)	(0.1490)
N	56	56	56	56	56	56
R Sq.	0.86	0.9	0.88	0.93	0.9	0.05

	(7)	(8)	(9)	(10)	(11)	(12)
	p5 (m+2 month)	**p5 (m+3 month)**	**p5 (m+4 month)**	**p5 (m+5 month)**	**p5 (m+6 month)**	**p5 upto Dec 2007**
Alpha	0.003439*	-0.001757	-0.003104	0.0006833	0.0001818	0.0053**
	(0.0020)	(0.0019)	(0.0020)	(0.0023)	(0.0018)	(0.0022)
Rm-Rf	0.8377***	0.8556***	0.7919***	0.8742***	0.7956***	0.84***
	(0.0864)	(0.0560)	(0.0599)	(0.0798)	(0.0479)	(0.0912)
SMB	-0.1921	0.04379	0.03651	-0.03306	-0.05414	0.0964
	(0.1278)	(0.1163)	(0.0995)	(0.1302)	(0.1053)	(0.1014)
HML	0.01371	0.04176	-0.1034	-0.07459	0.09220	-0.193*
	(0.1244)	(0.1022)	(0.0972)	(0.1177)	(0.0697)	(0.1103)
Momentum	-0.08947	-0.02224	-0.08002	0.09191	-0.01891	-0.185**
	(0.0610)	(0.0375)	(0.0483)	(0.0638)	(0.0332)	(0.0743)
N	56	56	56	56	56	44
R Sq.	0.89	0.86	0.87	0.84	0.82	0.78

Table I-6. Summary Statistics for Primary Variables

The following table shows the summary statistics of the variables used in the earnings regressions that follow this table. SVI is the average of the one month search index prior to the earnings announcement. NewsVol is the number of news stories about a firm that appear in the Google news archive search in the one month period prior to the earnings announcement. Assets are the total assets of the firm in millions. CAR5day is five day CAR centered on the earnings date. CAR3day is three day CAR centered on the earnings date. CAR30daypre is calculated from day -35 to day-5, where day 0 is the day of earnings announcement. The sample period is from January 2004 to December 2008.

	(1)	(2)	(3)	(4)	(5)	(6)
Percentile	**SVI**	**NewsVol**	**Assets**	**CAR5day**	**CAR3day**	**CAR30daypre**
1%	0.0000	0.00	123	-0.2425	-0.2283	-0.2914
5%	0.0000	0.00	303	-0.1308	-0.1175	-0.1449
10%	0.0000	0.25	508	-0.0866	-0.0786	-0.0965
25%	0.0000	0.67	1,113	-0.0347	-0.0307	-0.0418
50%	0.1217	1.75	3,042	0.0056	0.0049	0.0031
75%	0.2638	4.75	10,310	0.0503	0.0454	0.0496
90%	1.1051	14.75	31,069	0.1010	0.0947	0.1004
95%	1.7055	28.25	57,869	0.1407	0.1310	0.1432
99%	2.0782	84.50	287,583	0.2256	0.2168	0.2483
Mean	0.3078	6.29	19,843	0.0062	0.0061	0.0017
Variance	0.5093	14.99	91,577	0.0865	0.0798	0.0951
Min	0.0000	0	30	-0.7337	-0.6647	-0.9207
Max	2.3366	148	1,817,943	0.7577	0.6491	0.7443
Number of obs	15,345	15,345	15,345	15,345	15,345	15,345

Table I-7. Returns Around the Earnings Announcements

Specifications 1, 2 & 3 show the results of quarterly regression of the five-day cumulative abnormal return, centered on the earnings announcement date, on the lagged month over month change in SVI. Abnormal return is calculated by subtracting the value weighted market return from the stock return. SVIChange is the difference between average search index from day -35 to -5 and the average search index from day -66 to -36, where day 0 is the day of earnings announcement. NewsVolume is the number of news stories that mention the firm in the (-36 to -66) time period. SUE is standardized unexpected earnings defined as actual EPS less mean of analyst forecasts normalized by the standard deviation of analyst forecasts (Forecast Stdev). TwoMonthLagReturn is the return in the period from (-60,-120).

Following variables are based on the previous quarter's numbers: Log Size is natural log of sales. MarketToBook ratio is defined as the sum of total assets and market value of equity minus the book value of equity divided by total assets. Leverage is defined as long-term debt plus current liabilities divided by long-term debt plus current liabilities plus book value of common equity. ReturnOnAssets is income before extraordinary items divided by lagged assets value. SG&ARatio is the SG&A expense divided by sales.

For specifications 7 and 8, CAR is calculated from day (+5 to +30) and day (+5 to +45), where day 0 is the day of the earnings announcement. All days are calendar days. Standard errors in parentheses are White standard errors clustered by firm. The sample period is from January 2004 to December 2008. . ***, **, * denote 1%, 5% and 10% significance levels, respectively.

	Five Day CAR			Three Day CAR			Post 30day CAR	Post 45day CAR
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SVIChange$_{t-1}$	0.007224**	0.007081**	0.008036**	0.007302**	0.007096**	0.007777**	0.001179	0.0007817
	(0.0034)	(0.0034)	(0.0035)	(0.0031)	(0.0032)	(0.0032)	(0.0037)	(0.0043)
LogSize$_{t-1}$	-0.02434***	-0.02328***	-0.02397***	-0.02323***	-0.02239***	-0.02447***	-0.01983***	-0.03307***
	(0.0055)	(0.0055)	(0.0057)	(0.0055)	(0.0056)	(0.0057)	(0.0064)	(0.0071)
MarketToBookRatio$_{t-1}$	-0.01473***	-0.01492***	-0.01533***	-0.01218***	-0.01251***	-0.01241***	-0.01836***	-0.02171***
	(0.0024)	(0.0024)	(0.0026)	(0.0021)	(0.0021)	(0.0022)	(0.0021)	(0.0027)
Leverage$_{t-1}$	0.02450**	0.02779***	0.02448**	0.01721	0.02043*	0.02015*	0.004620	0.01035
	(0.0107)	(0.0106)	(0.0111)	(0.0105)	(0.0105)	(0.0107)	(0.0158)	(0.0184)
ReturnOnAssets$_{t-1}$	0.02375	0.01953	0.04453	0.02275	0.02239	0.05739	0.006758	-0.05318
	(0.0897)	(0.0889)	(0.0927)	(0.0926)	(0.0895)	(0.0925)	(0.0863)	(0.1099)
SG&A Ratio$_{t-1}$	0.007086	0.005671	-0.01582	0.0001457	-0.002441	-0.02370	-0.09547***	-0.1145***
	(0.0231)	(0.0229)	(0.0258)	(0.0219)	(0.0226)	(0.0249)	(0.0335)	(0.0358)
TwoMonthLagReturn	-0.02287***	-0.02325***	-0.02375***	-0.02041***	-0.02053***	-0.02107***	-0.01132*	-0.02356***
	(0.0073)	(0.0075)	(0.0078)	(0.0072)	(0.0074)	(0.0077)	(0.0066)	(0.0082)
SUE		0.001709***	0.002171**		0.001474**	0.001853*	0.0001222	-0.0002371
		(0.0007)	(0.0010)		(0.0006)	(0.0010)	(0.0001)	(0.0002)
Forecast Stdev		0.0005702	0.0005620		0.005790	0.005815	-0.02356	0.02577
		(0.0195)	(0.0203)		(0.0197)	(0.0204)	(0.0192)	(0.0261)
NewsVolume$_{t-1}$			0.00008158***			0.00007982***	-0.00003448	-0.00003825
			(0.0000)			(0.0000)	(0.0000)	(0.0000)
Constant	0.2058***	0.1978***	0.2090***	0.1920***	0.1858***	0.2035***	0.2029***	0.3099***
	(0.0391)	(0.0400)	(0.0416)	(0.0389)	(0.0399)	(0.0414)	(0.0462)	(0.0503)
Observations	11767	11340	10792	11767	11340	10792	10792	10792
Adjusted R-squared	0.0174	0.0440	0.0542	0.0149	0.0384	0.0471	0.0228	0.0344
Number of Firms	770	766	729	770	766	729	729	729
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table I-8. Returns Around the Earnings Announcements by Size

The following table reports the results of the regressions run on small and large firms separately. The sample is divided into three bins based on size. "Small Firms" constitute the firms in the smallest size bin while the "Large Firms" constitute the firms in the largest size bin. Specifications 1 and 2 show the results of regression of the five-day cumulative abnormal return, centered on the earnings announcement date, on the lagged Search Index. Abnormal return is calculated by subtracting the value weighted market return from the stock return. SVIChange is the difference between average search index from day -35 to -5 and the average search index from day -66 to -36, where day 0 is the day of earnings announcement. NewsVolume is the number of news stories that mention the firm in the (-36 to -66) time period. SUE is standardized unexpected earnings defined as actual EPS less mean of analyst forecasts normalized by the standard deviation of analyst forecasts (Forecast Stdev). TwoMonthLagReturn is the return in the period from (-60,-120).

Following variables are based on the previous quarter's numbers: Log Size is natural log of sales. MarketToBook ratio is defined as the sum of total assets and market value of equity minus the book value of equity divided by total assets. Leverage is defined as long-term debt plus current liabilities divided by long-term debt plus current liabilities plus book value of common equity. ReturnOnAssets is income before extraordinary items divided by lagged assets value. SG&ARatio is the SG&A expense divided by sales.

For specifications 5 and 6, CAR is calculated from day (+5 to +30) and for specifications 7 and 8, CAR is calculated from day (+5 to +45), where day 0 is the day of earnings announcement. All days are calendar days. Standard errors in parentheses are White standard errors clustered by firm. The sample period is from January 2004 to December 2008. ***, **, * denote 1%, 5% and 10% significance levels, respectively.

	Five Day CAR		Three Day CAR		Post 30 CAR		Post 45 CAR	
	Small Firms	Large Firms	Small Firms	Large Firms	Small Firms	Large Firms	Small Firms	Large Firms
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SVIChange$_{t-1}$	0.02051***	0.002528	0.01873***	0.001499	0.01224*	-0.006525	0.01878**	-0.006502
	(0.0072)	(0.0057)	(0.0072)	(0.0052)	(0.0070)	(0.0064)	(0.0086)	(0.0074)
LogSize$_{t-1}$	-0.03447***	-0.01140	-0.03646***	-0.008361	-0.03538***	-0.01078	-0.04902***	-0.01876
	(0.0100)	(0.0086)	(0.0106)	(0.0086)	(0.0114)	(0.0134)	(0.0132)	(0.0141)
MarketToBookRatio$_{t-1}$	-0.01549***	-0.02605***	-0.01308***	-0.02075***	-0.01577***	-0.01746***	-0.01661***	-0.02639***
	(0.0038)	(0.0058)	(0.0034)	(0.0057)	(0.0030)	(0.0056)	(0.0038)	(0.0055)
Leverage$_{t-1}$	0.01480	0.03038	0.01000	0.02309	-0.01121	0.04720*	-0.0008624	0.04766
	(0.0206)	(0.0206)	(0.0196)	(0.0203)	(0.0296)	(0.0241)	(0.0362)	(0.0295)
ReturnOnAssets$_{t-1}$	-0.01514	0.09375	0.006347	0.01537	0.04058	-0.1874	-0.1436	-0.3702
	(0.1851)	(0.1471)	(0.1812)	(0.1643)	(0.1369)	(0.1827)	(0.1690)	(0.2283)
SG&A Ratio$_{t-1}$	-0.006126	-0.1218**	-0.02189	-0.1134**	-0.1402***	-0.04715	-0.1348***	-0.1143
	(0.0391)	(0.0592)	(0.0395)	(0.0560)	(0.0460)	(0.0834)	(0.0476)	(0.0999)
TwoMonthLagReturn	-0.01248	-0.04076***	-0.009664	-0.03534**	-0.01911**	-0.01445	-0.02566**	-0.03839**
	(0.0120)	(0.0154)	(0.0128)	(0.0143)	(0.0089)	(0.0153)	(0.0111)	(0.0168)
SUE	0.001483**	0.006346***	0.001230*	0.005730***	-0.00001437	0.0008204*	-0.0003286**	-0.0002516
	(0.0006)	(0.0010)	(0.0007)	(0.0008)	(0.0001)	(0.0005)	(0.0001)	(0.0007)
Forecast Stdev	-0.01517	0.1219***	-0.02980	0.1456***	0.03860	-0.05089	0.1259	-0.01209
	(0.0526)	(0.0305)	(0.0536)	(0.0510)	(0.0989)	(0.0905)	(0.1612)	(0.0730)
NewsVolume$_{t-1}$	0.00008785***	0.00006897	0.00009106***	0.00008929**	-0.00005429	-0.00006075	-0.00001721	-0.00004096
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0000)	(0.0001)
Constant	0.2254***	0.1873**	0.2276***	0.1465*	0.2845***	0.1303	0.3635***	0.2533**
	(0.0591)	(0.0764)	(0.0620)	(0.0755)	(0.0659)	(0.1188)	(0.0739)	(0.1188)
Observations	3807	3409	3807	3409	3807	3409	3807	3409
Adjusted R-squared	0.0587	0.1104	0.0528	0.0992	0.0356	0.0201	0.0448	0.0358
Number of Firms	446	348	446	348	446	348	446	348
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table I-9. Pre-Earnings Announcement Returns

The following table reports the results of the regressions of 45 day and 30 day cumulative abnormal return on the lagged Search Index (SVI). Abnormal returns are calculated by subtracting the value weighted market return from the stock return. For specifications 1 to 3, the CAR is calculated from day -45 to day-5, where day 0 is the day of earnings announcement. For specifications 1 to 3, the SVIChange is the difference between average search index from day -46 to -76 and the average search index from day -76 to -106, where day 0 is the day of earnings announcement. For specifications 4 to 6, the CAR is calculated from day -35 to day-5. For specifications 4 to 6, the SVIChange is the difference between average search index from day -36 to -66 and the average search index from day -66 to -96. NewsVolume is the number of news stories that mention the firm in the prior one month time period. TwoMonthLagReturn is the return in the period from (-60,-120).

Following variables are based on the previous quarter's numbers: Log Size is natural log of sales. MarketToBook ratio is defined as the sum of total assets and market value of equity minus the book value of equity divided by total assets. Leverage is defined as long-term debt plus current liabilities divided by long-term debt plus current liabilities plus book value of common equity. ReturnOnAssets is income before extraordinary items divided by lagged assets value. SG&ARatio is the SG&A expense divided by sales.
All days are calendar days. Standard errors in parentheses are White standard errors clustered by firm. The sample period is from January 2004 to December 2008. ***, **, * denote 1%, 5% and 10% significance levels, respectively.

	CAR Pre 45 days			**CAR Pre 30 days**		
	(1)	(2)	(3)	(4)	(5)	(6)
SVIChange$_{t-1}$	0.006441*	0.009448**	0.009944**	0.01008	0.01202	0.01103
	(0.0037)	(0.0040)	(0.0041)	(0.0067)	(0.0075)	(0.0075)
LogSize$_{t-1}$	-0.03781***	-0.04643***	-0.04737***	-0.02668***	-0.03255***	-0.03006***
	(0.0059)	(0.0069)	(0.0072)	(0.0051)	(0.0062)	(0.0065)
MarketToBookRatio$_{t-1}$	-0.01396***	-0.01604***	-0.01771***	-0.008580***	-0.01023***	-0.01102***
	(0.0024)	(0.0029)	(0.0031)	(0.0025)	(0.0032)	(0.0035)
Leverage$_{t-1}$	-0.004134	-0.004105	-0.002902	-0.003195	0.005601	0.009294
	(0.0143)	(0.0158)	(0.0166)	(0.0125)	(0.0139)	(0.0144)
ReturnOnAssets$_{t-1}$		-0.1650	-0.1517		-0.1296	-0.1104
		(0.1192)	(0.1207)		(0.0978)	(0.0989)
SG&A Ratio$_{t-1}$		-0.05338	-0.06660		-0.03507	-0.01656
		(0.0398)	(0.0456)		(0.0233)	(0.0317)
TwoMonthLagReturn		-0.01840*	-0.01942*		-0.01097	-0.01321
		(0.0096)	(0.0099)		(0.0092)	(0.0097)
NewsVolume$_{t-1}$			-0.000007982*			-0.000007972**
			(0.0000)			(0.0000)
Constant	0.2708***	0.3578***	0.3731***	0.1944***	0.2497***	0.2328***
	(0.0385)	(0.0510)	(0.0532)	(0.0335)	(0.0445)	(0.0477)
Observations	14038	11655	11056	14038	11655	11056
Adjusted R-squared	0.0204	0.0270	0.0292	0.0151	0.0210	0.0223
Number of Firms	913	775	737	913	775	737
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Table I-10. Volume Prior to the Earnings Announcement

The following table reports the results of the regressions of buys volume ratios on the one month lagged Search Index (SVI). Daily buys and sells are calculated using data from the NYSE's TAQ database. The trades for each stock are signed based on the Lee and Ready (1991) algorithm. For specifications 1 to 4, the buys and sells are considered from day -45 to day-5, where day 0 is the day of earnings announcement. For specifications 5 to 8, the buys and sells are calculated from day -35 to day-5. The dependent variable in specification (1) is defined as follows: (sum of daily buys from (-45 to -5) / {(sum of daily sells from (-45 to -5) + (sum of daily buys from (-45 to -5)}. For specification (2), the dependent variable is the (sum of dollar daily buys from (-45 to -5)/ {(sum of dollar daily sells from day (-45 to -5)} + (sum of dollar daily buys from (-45 to -5)}. In specifications (3) and (4), I use the average of the buys ratio for each day to mitigate the effect of intense buying or selling days. Specifications 5 to 8 are similar to specifications 1 to 4, however the time period is from (-35 to -5).
Search volume index (SVI) is the average of one month prior search index. NewsVolume is the number of news stories that mention the firm in the one month prior period. TwoMonthLagReturn is the return in the period from (-60,-120)

Following variables are based on the previous quarter's numbers: Log Size is natural log of sales. MarketToBook ratio is defined as the sum of total assets and market value of equity minus the book value of equity divided by total assets. Leverage is defined as long-term debt plus current liabilities divided by long-term debt plus current liabilities plus book value of common equity. ReturnOnAssets is income before extraordinary items divided by lagged assets value. SG&ARatio is the SG&A expense divided by sales. All days are calendar days. Standard errors in parentheses are White standard errors clustered by firm. The sample period is from January 2004 to December 2008 and includes only the firms listed in NYSE for whom the TAQ data is available. ***, **, * denote 1%, 5% and 10% significance levels, respectively.

	Ratio of Buys to Total Volume							
	45 Days Prior to Earnings Announcement				30 Days Prior to Earnings Announcement			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SVI_{t-1}	0.01667***	0.01670***	0.4702***	0.4705***	0.01621***	0.01628***	0.02767***	0.02769***
	(0.0015)	(0.0015)	(0.0391)	(0.0391)	(0.0015)	(0.0015)	(0.0023)	(0.0023)
$NewsVolume_{t-1}$	-0.00001039**	-0.00001040**	-0.0002347*	-0.0002345*	-0.00001221***	-0.00001241***	-0.00001250**	-0.00001248**
	(0.0000)	(0.0000)	(0.0001)	(0.0001)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$LogSize_{t-1}$	-0.05418***	-0.05424***	-1.5122***	-1.5120***	-0.05476***	-0.05482***	-0.08896***	-0.08896***
	(0.0038)	(0.0039)	(0.1074)	(0.1074)	(0.0041)	(0.0041)	(0.0063)	(0.0063)
$MarketToBookRatio_{t-1}$	0.002228	0.002229	0.07421*	0.07353*	0.002701*	0.002726*	0.004360*	0.004320*
	(0.0016)	(0.0016)	(0.0430)	(0.0430)	(0.0016)	(0.0016)	(0.0025)	(0.0025)
$Leverage_{t-1}$	0.01215	0.01249	0.3097	0.3131	0.01237	0.01277	0.01821	0.01841
	(0.0091)	(0.0091)	(0.2491)	(0.2489)	(0.0091)	(0.0091)	(0.0147)	(0.0146)
$ReturnOnAssets_{t-1}$	0.1368***	0.1366***	3.2656**	3.2493**	0.1384**	0.1380**	0.1923**	0.1913**
	(0.0500)	(0.0500)	(1.3892)	(1.3872)	(0.0539)	(0.0540)	(0.0817)	(0.0816)
$SG\&A\ Ratio_{t-1}$	-0.1017***	-0.1022***	-2.8441***	-2.8425***	-0.1061***	-0.1062***	-0.1674***	-0.1673***
	(0.0245)	(0.0244)	(0.6615)	(0.6608)	(0.0260)	(0.0260)	(0.0389)	(0.0389)
TwoMonthLagReturn	0.008646***	0.008474***	0.2951***	0.2934***	0.009025***	0.009068***	0.01736***	0.01726***
	(0.0031)	(0.0031)	(0.0855)	(0.0853)	(0.0031)	(0.0031)	(0.0050)	(0.0050)
Constant	0.8890***	0.8900***	25.292***	25.298***	0.8918***	0.8924***	1.4878***	1.4882***
	(0.0297)	(0.0297)	(0.8262)	(0.8256)	(0.0314)	(0.0316)	(0.0486)	(0.0486)
Observations	8023	8023	8022	8022	8022	8022	8021	8021
Adjusted R-squared	0.1593	0.1596	0.1797	0.1799	0.1307	0.1313	0.1797	0.1799
Number of Firms	561	561	561	561	561	561	561	561
Firm Fixed Effects	yes	yes	yes	yes	yes	yes	yes	yes

References

Ball, R., and E. Bartov. 1996. How Naive is the Stock Market's Use of Earnings Information?, *Journal of Accounting and Economics* 21: 319-337.

Barber, Brad M., and Terrance Odean, 2000, Trading is hazardous to your wealth: The common stock investment performance of individual investors, *Journal of Finance* 55, 773-806.

Barber, Brad M., and Terrance Odean, 2002, Online investors: Do the slow do first? *Review of Financial Studies* 15, 455-488.

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, 785-818.

Battalio, R.H., and R. R. Mendenhall, 2005. Earnings expectations, investor trade size, and anomalous returns around earnings announcements, *Journal of Financial Economics* 77: 289-319.

Bernard, V. L., and J. Thomas. 1989. Post-Earnings Announcement Drift: Delayed Price Response or Risk Premium, *Journal of Accounting Research* 27: 1-35.

Choi, H., and H. Varian, 2009, Predicting the Present with Google Trends. Technical Report, Google Inc.

Chordia, Tarun*,* Richard Roll*, and* Avanidhar Subrahmanyam*, 2002,* Order imbalance, liquidity, and market returns*, Journal of Financial Economics 65, 111–130.*

D. W. Diamond and R. E. Verrecchia. "Information Aggregation in a Noisy Rational Expectations Economy." *Journal of Financial Economics* 9 (September 1981), 221-35.

Da, Z., Engelberg, J. and Gao, P. (2009a) "In Search of Attention", AFA 2010 Atlanta Meetings Paper. Available at SSRN: http://ssrn.com/abstract=1364209

Doyle, J.T., R.J. Lundholm, and M.T. Soliman. 2005. The Extreme Future Stock Returns Following Extreme Earnings Surprises. Working Paper, Stanford University.

Drake, M., D. Roulstone, and J. Thornock. 2011. Investor information demand: Evidence from Google searches around earnings announcements. Working paper, The Ohio State University.

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

Fama, E., and K. R. French. 1996. Multifactor Explanations of Asset Pricing Anomalies, *Journal of Finance* 51: 55–84.

Foster, G., C. Olsen, and T. Shevlin. 1984. Earnings Releases, Anomalies, and the Behavior of Security Returns, *The Accounting Review*: 574-603.

Francis, J., K. Schipper, and L. Vincent. 2002. Expanded Disclosures and the Increased Usefulness of Earnings Announcements, *The Accounting Review* 77: 515-546.

Grossman, S. and J. Stiglitz. "On the impossibility of informationally efficient markets." *American Economic Review* 70 (1980): 393–408.

Gu. 2005. Innovation, Future Earnings, and Market Efficiency, *Journal of Accounting, Auditing & Finance* 20: 385-418.26.

Harris and Raviv (1993) M Harris and A Raviv, Differences of opinion make a horse race. *Review of Financial Studies*, 6 (1993), pp. 473–506.

Holthausen, R. and R. Verrecchia. "The effect of informedness and consensus on price and volume behavior." *Accounting Review* 65 (1990): 191–208.

Hyunyoung Choi, Hal Varian. 2009 (working paper). Predicting the Present with Google Trends Google Inc.

Imhoff, E.A. 1992. The Relation between Perceived Accounting Quality and Economic Characteristics of the Firm, *Journal of Accounting and Public Policy* 11: 97-118.

Kaniel, Ron, Gideon Saar, and Sheridan Titman, 2008, Individual investor trading and stock returns, *Journal of Finance* 63, 273–310.

Karpoff, J., 1987, The Relation between Price Change and Trading Volume: A Survey, *Journal of Financial and Quantitative* Analysis, 22, 109-126.

Kim, O. and R. Verrecchia. "Pre-announcement and event-period private information." *Journal of Accounting and Economics* 24 (1997): 395–419.

Kothari, S.P., 2001. Capital markets research in accounting. *Journal of Accounting and Economics* 31, 105–231

Lee, C. and M. Ready. "Inferring trade direction from intraday data." *Journal of Finance* 46 (1991): 733-746.

Lev, B. 1989. On the Usefulness of Earnings and Earnings Research: Lessons and Directions from Two Decades of Empirical Research, *Journal of Accounting Research* 27: 153-192

Liu, J., and Thomas, J. 2000. Stock Returns and Accounting Earnings, *Journal of Accounting Research* 38: 71-101

Livnat, J., and R.R. Mendenhall. 2006. Comparing the Post-Earnings Announcement Drift forSurprises Calculated from Analyst and Time-Series Forecasts, *Journal of Accounting Research* 44: 177-204.

Schwert, 1981 G.William Schwert, The adjustment of stock prices to information about inflation. *Journal of Finance*, 36 (1981), pp. 15–29.

Chapter II

Effect of Employee Satisfaction on Earnings Surprises

Abstract

Using a unique data set drawn from self administered employee surveys for 1495 US public corporations; I construct an Employee Satisfaction Index (ESI) and use it as a proxy for employee satisfaction. I find that ESI is higher for larger firms, high market to book ratio firms and firms that have low leverage. I also look at the effect of the changes in ESI on quarterly earnings announcements. I find that the changes in ESI are positively and (weakly) significantly related to the future quarterly earnings surprises. Moreover, the effect is stronger for companies that have higher information asymmetries and are more human capital dependent. The results are consistent with the theories that state that employees are insiders in a company and have information relevant to the future corporate performance. Moreover, consistent with human-capital centric theories I find evidence that the change in employee satisfaction has a greater effect on the performance of human-capital dependent companies.

2.1 Introduction

Do changes in employee satisfaction influence future earnings surprises? Using a novel dataset based on self administered employee surveys, I find that the changes in employee satisfaction are positively and significantly related to the future quarterly earnings surprises. Moreover, the effect is stronger for companies that have higher information asymmetries and are more human capital dependent.

Most of the empirical work tying employee satisfaction to future performance of a company is limited to the Fortune magazine's list of the "100 Best Companies to Work for in America." The list is published every year in January and was first published in 1998. Filbeck and Preece (2003) analyze the effect of inclusion in the 1998 Fortune's Best Companies list on returns from 1987-1999. They find that companies in the Best Companies list do not outperform size- and industry-matched benchmarks.

However, recent studies have shown that employee satisfaction is positively related to the future performance of a company. Faleye and Trahan (2006) find that investors react positively to the announcement of inclusion of companies in the Best Companies list, and that list firms outperform a size-matched and industry-matched control group on productivity, profitability, and value creation. Edmans (2011), analyzes the relationship between employee satisfaction and long-run stock returns. He finds that portfolio of Best Companies earns a significant risk-adjusted alpha. He also finds that the companies in the Best Companies list had significantly more positive earnings surprises and announcement returns.

Above studies indicate that employee satisfaction has some information content, that is revealed to the market when the company is included in the Best Companies list. However, none of these studies examine if the change in employee satisfaction has information that can influence the outcome of future information events such as mergers, dividend announcements, IPO success and

earnings announcements. Primary reason for this deficiency is the absence of employee satisfaction data at a frequent level. Using a novel dataset based on self administered employee surveys, I overcome this limitation and study the information content of changes in employee satisfaction. Specifically, in this paper I analyze the relationship between changes in employee satisfaction and future earnings surprises.

Employees are insiders in a company who directly observe the effect of changes in products, services and policies taking place in the company. For example, employees who sell the new product or service know firsthand whether the product or service is successful. Working and being part of a successful launch boosts morale among the employees and gives them a sense of accomplishment. This may translate into a positive review of the company.

A priori it is not clear if the changes in employee satisfaction will have a positive or negative effect on future performance. If a company is making changes that lead to increase in employee satisfaction at the expense of shareholders, we can expect that the increase in employee satisfaction will lead to worst future performance. Abowd (1989) demonstrates that announcements of pay increases are associated with reduction in market value of a company. Gorton and Schmid (2004) find that greater employee involvement is associated with reduction in future profitability and valuation.

However, if a company makes investment in employees that motivate employees to perform better then we can expect that the increase in employee satisfaction will have a positive effect on future performance. McGregor (1960) argues that higher levels of employee satisfaction lead employees to identify with the company and its goals and thus induces effort. Akerlof and Yellen (1986), (Akerlof, 1982) posit an efficiency wage theory, under which increased employee satisfaction can increase effort, because an employee does not want to get fired from a satisfying job.

Alternatively, an employee reciprocates the "gift" of increased satisfaction with increased effort upon his part.

To test the above hypothesis, I use a dataset based on employee generated surveys for US public companies from January 2008 to May 2011. In each of the reviews, employees rate their company on seventeen variables such as career prospects, work-life balance and fairness and respect. Complete list of the variables is shown in appendix II-1. Employees rate their employer on a scale of 1 to 5. 1 represents that an employee is highly dissatisfied while 5 represents she is highly satisfied. To determine the main sources of variation in the seventeen dimensional dataset, I run principal component analysis and find that one principal component explains almost 60% of the variation within the data. The first principal component is the weighted average of each of the variables, with weights being almost equally across all the variables. This component can be interpreted as the overall satisfaction of an employee with the company. Based on the first principal component, I construct an employee satisfaction index (ESI) by adding responses to each of the variables.

First, I check if the ESI is merely proxying for characteristics of a firm. I regress the three month average ESI index on the characteristics of a firm drawn from the quarterly announcements data from Compustat. I find that ESI tends to be higher for companies that have a high market to book ratio, are large and have low leverage. However, low R square of less than five percent indicates that most of the variation in ESI is not being explained by the firm characteristics.

To explore the information content of ESI, I look at information events, such as quarterly earnings announcements, and analyze if the past changes in ESI can predict quarterly earnings surprises. Earnings surprise is measured by the three-day or five-day cumulative abnormal return around the earnings announcements. I find that for the overall sample changes in ESI weakly (significant at five-percent level) predicts future earnings surprises. Consistent with the information story, I find that the effect of ESI on future earnings surprises is stronger for firms where

information asymmetries are higher. Results are also stronger when the quarters most affected by the 2008 financial crisis are excluded.

As a further support of the information hypothesis, I run the analysis for companies where human capital is deemed important. Employees of companies where human capital is relatively material will have greater influence, and their morale and satisfaction will have a greater impact on future performance. Moreover, the reviews by these talented employees will have a higher informational content. I find that firms where human capital is deemed to be important, the results are much stronger both in terms of significance and magnitude of the coefficients.

I also consider three alternative hypotheses that may explain the results. First, it is possible that the ESI is merely capturing characteristics of the firm such as size or market to book ratio. However, the results are robust to the inclusion of firm characteristics. Second, it is possible that ESI is merely reflecting the information contained in analyst reports. I control for changes in analyst forecast and information environment surrounding the earnings surprises and find that results are still significant. Third, it is possible that ESI is merely reflecting the broader investor satisfaction index (Dorn, Huberman, and Sengmueller (2008) and Barber, Odean, and Zhu (2009)). Results are robust to the inclusion of time fixed effects as well as past returns.

This paper is organized as follows: section 2 describes the data and salient issues related with the data used in the paper; section 3 presents the results and section 4 summarizes and concludes.

2.2 Data

2.2.1 Source of Data

Data for the project was provided by Glassdoor.Com (GD). GD provides a forum where employees can rate the place where they work at. The website has been profiled by leading business

magazines such as BusinessWeek and Forbes. GD's founders had previously founded successful startups such as Hotwire.com and Expedia.com. According to their website-

"Glassdoor is a free jobs and career community that offers the world an inside look at jobs and companies. What sets us apart is our "employee generated content" – anonymous salaries, company reviews, interview questions, and more – all posted by employees, job seekers, and sometimes the companies themselves. Now with nearly 3 million salaries and reviews, you have all the information you might need to make your next career decision."

In present paper, I focus on the employee reviews of the companies. An employee can log into the GD's website and rate her employer. A typical review on the website consists of a numerical review and a written review. An employee rates his employer based on seventeen dimensions such as work-life balance, fairness and respect within the company, and information and knowledge sharing. The complete list of seventeen questions appears in appendix II-1. As can be seen from the appendix, there are primarily nine dimensions that are being covered by the seventeen survey questions. For example, following two questions are asking about employee morale -

Question 1. Company XYZ as a place you would recommend to others to work?

Question 2. Company XYZ as a place you are proud to work?

Employees rate their employer on a scale of 1 to 5. A rating of one implies the employee is very dissatisfied while a rating of five implies employee is very satisfied. Employees also write a paragraph or two about the company. In present paper, I restrict myself to the numerical reviews.

2.2.2 Data Description

Table II-1 shows the snapshot of dataset that is used for the present paper. Each line represents a single review. Each review consists of the details about the company, date of the review

and ratings entered by an employee for each of the seventeen variables. Employees rate the company on a specific dimension by assigning a numerical value from 1 to 5. A rating of 1 implies that an employee is very dissatisfied with that particular aspect of the company. A rating of 5 implies very satisfied, with 3 being neutral. In addition to rating a company, employees also enter their occupation and work location. Additionally, one can look at the number of people who found the review as helpful as well as see if an employee is a current or former employee.

2.2.3 Sample Description

Current sample is from January 2008 to May 2011. Jan 2008 is the first month the data was available. While May 2011 is the last month, the data was pulled from the website for the present paper. While employees rate their company on a number scale as well as write a review of the company. I restrict myself to the numerical reviews. It is possible to glean additional information by doing keyword searches on the written part of the review. Reviews on the GD website constitute companies that are private or public, US based or Non-US based. I restrict my sample to the US public firms since it is easy to find supplemental data from standard data sources such as CRSP and Compustat for these firms. This leaves me with a sample of 71,567 reviews.

Since I look at changes in the reviews prior to the earnings announcements, I drop all firms that have less than two reviews in the whole sample period from January 2008 to May 2011. Final sample consists of 1495 US public firms. Table II-2 shows the distribution of reviews by SIC codes. Companies in the SIC code related to manufacturing constitute the largest fraction of the sample, while companies in the food industry constitute the smallest fraction.

As shown in Table II-3a number of reviews per firm vary substantially. The median number of reviews per firm is 147 and standard deviation is 374. In general larger firms that have a larger number of employees tend to get reviewed more often than smaller firms. Table II-3b shows the

results of regression of the total number of reviews of a firm on firm characteristics for the first quarter of 2011. Clearly, size of a firm has a significant effect on the number of reviews a firm has. To explore the effect of industry on reviews, I add dummy variables for SIC codes. As shown in the Table II-3 industry dummies do not have any significant coefficients that may explain the number of reviews a firm gets.

2.2.4 Variable Description

As shown in Table II-4 employees rate their employers on 17 variables. A closer look at the questions shows that there are underlying nine dimensions in these 17 variables. The dimensions are:

1) Career opportunities

2) Compensation and benefits

3) Communication

4) Employee morale

5) Feedback and recognition

6) Senior management

7) Work life balance

8) Fairness and respect

9) Overall satisfaction with the company

For example, the following two questions are asking about recognition and feedback-

Question 1. Feedback you receive about your job performance?

Question 2. Recognition and praise you receive when you do a good job?

Since both questions are asking about the same underlying dimension (recognition and feedback), we can expect that the responses to the two questions will not be far apart. I utilize this aspect of

redundancy to weed out reviews that are being entered randomly or nonsensically. For example, it will be odd if an employee rated the two feedback-related questions widely differently. If I find that an employee has rated a company vastly differently (defined as a difference of 3 or more) on two or more questions, I exclude that review. This leads to exclusion of 80 reviews in the final sample of US public firms.

In addition, employees also give an overall rating to the firm. To check for any systematic bias in the reviews, I look at the average value of reviews along each of the dimensions. Table II-5 shows the average value and standard deviation of the ratings along the eight dimensions. While some dimensions vary from the unbiased value of three, the relatively high standard deviation makes these variations insignificant. On average, the reviews do not seem to be biased negatively or positively.

2.2.5 Principal Component Analysis

To further reduce the number of dimensions and to look at the leading sources of variation in the data, I run principal component analysis (PCA) on the sample. Using PCA one can get a better understanding of the main sources of variation in a sample with many dimensions. PCA helps to analyze data that consists of various variables that are also inter-correlated. PCA helps to extract important dimensions from the data and expresses them in variables that are orthogonal to each other. These dimensions or variables are called principal components.

Table II-6 shows the results of the PCA analysis. Three main components explain over seventy percent of the variation in the sample; with the first component explaining about 58% of the variation. First principal component weighs almost equally on each of the dimensions and can be interpreted as representing overall satisfaction of an employee with the company. Second principal component loads heavily on the compensation and benefits variables while the third principal

component loads heavily on the work-life balance variables. Interpretation of other principal components is not as straight forward.

2.2.6 Employee Satisfaction Index

2.2.6.1 Construction

In present paper, I focus on the first principal component as my main explanatory variable. To simplify the construction of the variable, I sum the ratings along each of the dimensions and call this variable employee satisfaction index. Table II-5 shows the summary statistics for this variable. The variable seems to be biased towards the positive side; however, high value of standard deviation makes the bias statistically insignificant.

2.2.6.2 Cross sectional Determinants of ESI

Since by construction the employee satisfaction index captures the overall satisfaction an employee feels for a company. It is worth looking at the firm characteristics that influence employee satisfaction. In Table II-7, I empirically test if common firm characteristics affect the employee happiness. Dependent variable is log of the average of three-month employee satisfaction index measured after the quarterly earnings announcement date. Firm characteristic variables are measured at the quarterly level at the time of the earnings announcement. Following equation is evaluated in specification one in Table II-7.

$$LogESI_{(i,t)} = \alpha + \beta_1 \text{LogAssets}_{i,t-1} + \beta_2 \text{MktToBook}_{i,t-1} + \beta_3 \text{LeverageMkt}_{i,t-1} + \beta_4 \text{ReturnOnAssets}_{i,t-1} + \beta_5 \text{SG\&A}_{i,t} + \varepsilon_{i,t-1} \quad (EQ1)$$

Number of firms i =1, 2,… N

Number of quarterly earnings announcements $t = 1, 2, \ldots N_q$

In specification (2), industry dummies are added to the analysis. None of the industry dummies are significant and are not shown in the table. Specification (3) excludes all financial services firms. Specification (4) uses two-stage Fama-Macbeth approach. In each of these specifications, employee satisfaction index is significantly higher for firms that are larger, have high value of market to book ratio and have lower leverage. Larger firms may provide more job security, recognition and opportunities to move around. Firms with high market to book ratio are most likely firms where human capital is more important and hence not only does the kind of work that employees do is more challenging but also the firms may be making extra efforts to keep the high-value employees. Higher leverage can be associated with financial distress and decrease in job security, and this may have an impact on employee satisfaction. To ensure that the high leverage financial firms are not having undue influence on the results, I rerun the analysis after excluding the firms in the SIC codes 6000-6999. The results are robust to the exclusion of these firms.

2.3 Empirical Results

2.3.1 Hypothesis Development

Earnings announcement periods provide the management of the company with an opportunity to aggregate the operating and financial results and share it with the public. Earnings surprise indicates that the company performed better or worse than the market expectations. Earnings surprise is the new information that while present in disaggregate form among the diverse employees and customers of the company, is revealed to the market only when the information has been aggregated by the management of the company and shared with the public on the day of the earnings announcement.

Changes in the ESI also mirror the above aggregation of information. A multitude of employees of the company observe the company's products and services first hand. Employees being insider to the company observe the effect of change in policies, service or launch of new products firsthand. For example, employees who sell the new product or service know firsthand whether the product or service is experiencing success. Working and being part of a successful launch boosts morale of the employees and gives them a sense of accomplishment. This may translate into a positive review of the company they work for. If that is the case, we should expect that an uptick in the employee reviews of the company to have information content relevant to the future earnings surprises. The effect is also expected to be stronger for companies that have high information assymteries.

I also consider alternative channels through which employee satisfaction index may affect future returns of the company. A priori it is not clear if the changes in employee satisfaction will have a positive or negative effect on future performance. If a company is making changes that lead to increase in employee satisfaction at the expense of shareholders, we can expect that the increase in employee satisfaction will lead to worst future performance.

However, a higher value of ESI can imply that a company takes good care of its employees. This in turn attracts better employees and increases the retention rate of star employees. Thus, a company develops a pool of valuable human capital that translates into better future performance. In such a case, we should expect that superior performance continues in future and is not just limited to the current earnings period.

Since human capital dependent firms tend to employ skilled and well-educated workers, we can expect that their reviews will be more informational. Moreover, since they have greater influence on the performance of the company changes in their satisfaction levels should have a greater impact

on performance. Hence, we should expect that the effect of ESI on earnings will be greater for firms where human capital is more important.

2.3.2 Choice of Event - Quarterly Earnings Announcements

To test the above hypothesis, I consider events where information is dispensed to the public. The empirical question that is addressed using this method is to check whether the ESI can predict the outcome of the announcement. Possible candidates for such events are merger announcements, dividend changes announcements and earnings announcements. I chose earnings announcements as the event to study in this paper. There are various advantages to using earnings announcements. Mergers and dividend announcements suffer from potential selection bias issues since companies making these announcements are not randomly distributed. All public firms are obligated by law to announce earnings every quarter. Additionally, this also leads to more observations since I have four data points for each firm each year. Moreover, there is a rich information environment surrounding the earnings announcements such as analyst reports that lets me proxy for market expectations.

2.3.3 CAR as a measure of earnings surprises

Earnings surprise occurs when the actual numbers are better or worse than what the market was expecting. Market expectations can be proxied in various ways. I will consider the two most common methods. First method uses analyst forecasts as a proxy for market expectations. For example, median or mean earnings per share number is calculated using the I/B/E/S dataset prior to the earnings announcements. Actual EPS is subtracted from the median or mean forecasted number and divided by the standard deviation of the analysts forecasts to come up with Standardized Unexpected Earnings (SUE). Standard deviation of analyst forecasts for that quarter proxies for the information uncertainty surrounding the earnings announcements.

One can also measure an earnings surprise by looking at CAR centered on the earnings. There are various advantages to using CAR as a measure of earnings surprise. First, CAR does not suffer from the measurement error that SUE suffers from since for SUE one has to estimate the market's unobservable expectations of earnings. Moreover, during the earnings announcement management discloses other information apart from earnings. Management provides extensive disclosure through financial reports, footnotes, management discussion of the results, forecasts and other forward-looking information.

In addition, many firms engage in conference calls where top executives of the firm present the last quarter's results and answer questions from analysts and investors. It is very difficult to capture the full information content of earning's surprise by looking at SUE alone. This is consistent with the findings of Francis, Schipper and Vincent (2002) who find that earnings announcements increasingly serve as a conduit for information different from merely the earnings per share numbers. In present paper, I use CAR centered on earnings announcement date as a measure of earnings surprise.

2.3.3.1 Quarterly Earnings Announcements Date

Publicly owned companies are required by law to file quarterly earnings reports with the SEC. Many companies announce the earnings announcement dates in advance, but several companies miss these dates and report late. To determine the actual date of earnings announcements, I use Compustat. Compustat records the earnings date as the date when the earnings report appears in the Wall Street Journal or other newspapers. However, the coverage of the Wall Street Journal is biased towards well known large stocks, and it is possible that earnings announcement dates of smaller stocks are not accurate. To be certain that I have the correct earnings announcement dates, I cross check the earnings dates for smaller stocks against the Factiva

database. I check for news stories to confirm that the date in Compustat is the actual earnings announcement date. In case of conflict, I use the date provided by the news story in Factiva.

2.3.4 Results

2.3.4.1 Effect of changes in ESI on earnings surprises

CAR centered on earnings announcement date is regressed on Log ESI and control variables. General specification for the regression equation is -

$$CAR_{(i,t-2,t+2)} = \alpha_{1i} + \alpha_{2t} + \beta_1 \text{LogESI}_{i,t-90} + \beta_2 \text{LogAssets}_{i,t-1} + \beta_3 \text{MkttoBook}_{i,t-1} + \beta_4 \text{LeverageMkt}_{i,t-1} + \beta_5 \text{ReturnOnAssets}_{i,t} + \beta_6 \text{SG\&A}_{i,t} + + \beta_7 \text{AnalystStdev}_{i,t} + \varepsilon_{i,t} \quad (EQ2)$$

Number of firms i =1, 2,… N

Number of quarterly earnings announcements t =1, 2,…Nq

CAR is five days cumulative abnormal return centered on the earnings announcement date. Abnormal return is defined as the return minus the value weighted market return. ESI is calculated by averaging the reviews in the 90-day window prior to the earnings announcement date and taking the log of the average. Since reviews are randomly distributed over time and are not periodic it is important to have a large enough window so that one can draw upon a substantial number of reviews before each earnings date. 90-day window is the maximum duration window that can be selected between the quarterly earnings announcement dates. In the 90-day window, I exclude 10days prior to the announcement date to take into account any leakage of earnings information. As a robustness test, I also exclude ten days after the previous earnings announcement date, to account for any immediate effect that previous earnings announcement may have on the ESI.

Lagged values of Size, Book to Market Ratio and Leverage variables are added to the equation to control for their effect on returns and to account for the fact that ESI is significantly related with these firm characteristics. Lagged value of profitability measure controls for the effect of firm profitability on employee satisfaction. Lagged value of SG&A expense divided by total assets proxies for effect of advertising and marketing on employees that may influence their reviews. Standard deviation of analyst's forecasts is used as a proxy for the uncertainty surrounding the earnings announcement period.

Specification in Table II-8 shows the results of regressing five-day CAR (from day-2 to day +2, where day zero is the day of the earnings announcement) on ESI and control variables. Specification (1) and all specifications that appear after that use firm fixed effects. Standard errors clustered by firm are shown in parenthesis. As shown in Table II-8 specification (1) a 10% increase in ESI index is related with a 14 basis points of excess return. Specification (2) includes quarterly time fixed effects to account for any observable or unobservable variables that are constant over time. Results hold both in magnitude and significance with this specification.

Specification (3) includes one month return 30 days prior to the earnings announcement and two-month return 60 days prior to the earnings announcement. These are included to control for momentum effect and to check if the results are being driven by past extreme returns. Results are robust to inclusion of these two variables.

2.3.4.2 Financial crisis of 2008

Sample period from January 2008 to May 2011 was marred by the financial crisis in late 2008. To mitigate the effect of the crisis on the results, I exclude last quarter of 2008 and first quarter of 2009. Table II-8, specification (4) shows the results of this regression. Once the most troubled quarters are taken out the results are stronger in terms of both significance and magnitude.

Specifically, the magnitude of the coefficient of ESI increases by almost 35% and is significant at 1% level in contrast to being significant at 5% level earlier. As a robustness check, I also confirm that I get similar results when I exclude the last quarter of 2008 and the first two quarters of 2009.

2.3.4.3 Results for companies where information asymmetries are higher

If indeed, ESI is capturing information pertaining to future performance of the company, the information should be more valuable for companies where information asymmetries are higher. In general, smaller companies have lower analyst following, less news coverage and are not major part of institutional investors' holdings. In specifications (5) and (6) in Table II-8, I test if ESI performs better for smaller companies versus larger companies. The full sample is divided into three bins based on size. "Small Firms" constitute the firms in the smallest size bin while the "Large Firms" constitute the firms in the largest size bin. Specification one is run for the firms in the smallest size bin, and specification (2) is run for firms in the largest size bin. Coefficient of ESI for smaller firms is significant at 1% level and is greater than the coefficient of the overall sample by over 60%. ESI is more strongly associated with future surprises for firms where information asymmetries are higher.

2.3.4.4 Results for companies where human capital is important

In current paper, I propose that the changes in ESI predicts future earnings surprises since employees reviews reflect the condition of the company from point of view of an insider. Hence we should expect that in companies where human capital is important the results will be stronger. Specifically, employees of such companies will be a more influential factor in production, and hence their morale and satisfaction will have a greater impact on future performance. Moreover, the reviews by these talented employees will have a higher informational content.

In the existing literature, there are no good proxies for human capital in a firm. For my purpose, I only need to proxy for companies where human capital is more important. I use three proxies for human capital:

a) Market to Book Ratio
b) R&D expense scaled by lagged total assets
c) SG&A expense scaled by lagged total assets

I argue that firms that have more intangible assets with respect to tangible assets are firms where human capital is potentially more important. As a simple measure, I use market to book ratio for the firms to proxy for the importance of human capital in a firm. Firms that have a higher market to book ratio tend to have a greater number of intangible assets per unit of tangible assets than firms with a lower market to book ratio. The full sample is divided into three bins based on the market to book (m/b) ratio. "High m/b" constitutes the firms in the largest market to book ratio bin while the "Low m/b" constitute the firms in the smallest market to book ratio bin. Specification (1) is run for the firms in the smallest market to book ratio bin and specification (2) is run for firms in the largest market to book ratio bin. Coefficient of ESI for high market to book ratio firms is significant at 1% level while it is insignificant for the low market to book ratio firms. As shown in Table II-9, I get similar results when I divide the sample based on R&D expense to assets and SG&A to sales. ESI is more strongly associated with future surprises for firms where R&D expense to assets and SG&A to assets are higher.

2.4 Conclusion

Using a novel dataset based on self-administered employee surveys, I construct an Employee Satisfaction Index (ESI) variable. I find that ESI is higher for larger firms, high market to book ratio firms and firms that have low leverage. I also look at the effect of the changes in ESI on future corporate events. In this paper, I focus exclusively on quarterly earnings announcements. I find that the changes in ESI are positively and significantly related to the future quarterly earnings surprises. Moreover, the effect is stronger for companies that have higher information asymmetries and are more human capital dependent. This is consistent with the theories that purport that employees are insiders in a company and have information relevant to the future corporate performance. Moreover, consistent with human-capital centric theories I find evidence that the change in employee satisfaction has a greater effect on the performance of human-capital dependent companies.

Table II-1. Employee Review Data

This table shows the snapshot of data that is used for the present paper. Each line represents a single review. Each review consists of the details about the company, date of the review and ratings entered by an employee for each of the seventeen variables. Employees rate the company on a particular dimension by assigning a numerical value from 1 to 5. Q1-Q17 correspond to questions detailed in appendix II-1.

EmployerName	City	Occupation	ReviewDate	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17
Walgreens	Dallas	Pharmacy Manager	9/1/2009	5	5	5	5	5	4	5	4	5	3	4	4	5	2	5	5	4
Sikorsky	Fort Worth	Design Engineer	9/1/2009	1	1	4	4	2	2	3	3	4	3	2	2	4	2	2	2	3
Marsh USA	New York	Vice President	9/1/2009	3	3	2	4	3	4	3	3	2	1	3	3	5	5	2	2	2
Fastenal	Burlington	General Manager	9/1/2009	3	3	3	2	4	1	1	2	2	1	1	1	3	2	1	2	1
VMware	Palo Alto	Consulting Architect	9/1/2009	2	2	2	3	2	1	2	2	1	3	2	2	4	3	2	1	1
3M Precision Optics	Cincinnati	NULL	9/1/2009	4	3	5	3	5	4	2	3	4	4	4	4	5	2	3	4	3
ICF International	Fairfax	Associate	9/1/2009	2	2	2	3	3	3	2	3	3	4	3	3	3	2	2	3	2
Shell Oil	Houston	NULL	9/1/2009	4	4	4	5	3	3	3	3	3	3	2	3	4	3	2	3	3

Table II-2. Distribution of Sample Firms

Following table lists the distribution of sample companies by SIC Codes. The sample period is from January 2008 to May 2011.

SIC Codes	Industry Discription	Number of Companies	Percent
0-1000	Food products	3	0.2%
1000-1999	Mining and construction	53	3.5%
2000-2999	Consumer products	200	13.4%
3000-3999	Manufacturing	348	23.3%
4000-4999	Utilities and transportation	141	9.4%
5000-5999	Wholesale, retail and some services	196	13.1%
6000-6999	Financial services	233	15.6%
7000-7999	Personal & business services	245	16.4%
8000-8999	Miscellaneous	76	5.1%
Total		1,495	100%

Table II-3a. Number of Reviews per Firm I

This table shows the summary statistics for the number of reviews per firm for the sample 1495 US public firms. The sample period is from January 2008 to May 2011.

Percentiles	Number of Reviews
1%	2
5%	6
10%	12
25%	44
50%	147
75%	455
90%	820
95%	1,175
99%	1,566
Mean	309.21
Std. Dev.	373.60

Table II-3b. Number of Reviews per Firm II

Following table shows the results of regression of total number of reviews per firm on firm characteristics for the first quarter of 2009. Industry dummies were used for each one digit SIC Code.

Number of Reviews by Firm	Coefficient	Standard Error
Food products	-34.43007	45.652
Mining and construction	-30.11084	44.554
Consumer products	14.34385	44.336
Manufacturing	15.65385	45.135
Utilities and transportation	51.95731	44.526
Wholesale, retail and some services	6.975378	44.882
Financial services	42.04352	44.527
Personal & business services	37.71163	45.299
Miscellaneous	-184.5965***	67.939
Log Assets	13.845***	1.958
Market to Book Ratio	0.964	2.198
Leverage (market)	-9.456	16.695
Return on Assets	2.386	122.119
SG&A over Assets	33.912***	12.391
Constant	-113.615**	46.664
Adjusted R-squared	0.3923	
Number of observations	1,265	

Table II-4. Primary Dimensions of Survey Questions

Following table matches the survey questions with the primary underlying dimensions.

Question Number	Survey Questions	Classified in the present paper as:
	How satisfied are you with:	
1	Your opportunities for professional growth at IBM?	Career Opportunities 1
2	Your opportunities for career advancement within IBM?	Career Opportunities 2
3	Your compensation compared with similar jobs elsewhere?	Compensation and Benefits 1
4	Your benefits package compared with similar employers?	Compensation and Benefits 2
5	Information and knowledge sharing within IBM?	Knowledge Sharing 1
6	Communications from management about important issues and changes?	Knowledge Sharing 2
7	IBM as a place you would recommend to others to work?	Employee Morale 1
8	IBM as a place you are proud to work?	Employee Morale 2
9	Feedback you receive about your job performance?	Feedback and Review 1
10	Recognition and praise you receive when you do a good job?	Feedback and Review 2
11	Leadership abilities of Senior management?	Sr. Leadership 1
12	Competence of Senior management?	Sr. Leadership 2
13	Management support in permitting time off when you think it's necessary?	Work Life Balance 1
14	Employer support in balancing between work life and personal life?	Work Life Balance 2
15	Fairness in how promotions are given and people are treated?	Fairness and Respect 1
16	The level of respect shown by management toward employees?	Fairness and Respect 2
17	Overall, how satisfied are you with IBM as a place to work?	Overall

Table II-5. Summary Statistics

This table shows the summary statistics for the primary variables that constitute the survey questions for the sample 1495 US public firms. The sample period is from January 2008 to May 2011. ESI is Employee Satisfaction Index and is sum of all the scores for each variable in the survey.

Variable	Observations	Mean	Std. Dev.	Min	Max
Overall	71,567	3.07	1.17	1	5
Work Life Balance	71,567	3.48	1.16	1	5
Career Opportunities	71,567	2.98	1.13	1	5
Compensation and Benefits	71,567	3.30	0.98	1	5
Knowledge Sharing	71,567	3.03	1.09	1	5
Employee Morale	71,567	3.13	1.22	1	5
Feedback and Review	71,567	3.11	1.14	1	5
Sr. Leadership	71,567	2.84	1.25	1	5
Fairness and Respect	71,567	2.85	1.16	1	5
ESI	71,567	52.55	15.67	17	85

Table II-6. Principal Component Analysis

Following table shows the results of the Principal Component Analysis run on the whole sample of 71,567 reviews for 1495 US public corporations. Loadings of variables on each component are shown in the table. Three main components explain over seventy percent of the variation within the sample; with the first component explaining about 58% of the variation. First principal component weighs almost equally on each of the dimensions and can be interpreted as representing overall satisfaction of an employee with the company. Second principal component loads heavily on the compensation and benefits variables while the third principal component loads heavily on the work-life balance variables. Based on the first principal component Employee Satisfaction Index is sum of all the survey scores for each variable in the survey.

Variable	Component 1	Component 2	Component 3	Component 4	Component 5
Overall	0.2934	0.0374	-0.019	-0.0051	-0.1199
Career Opportunities 1	0.2491	0.2114	-0.1722	-0.39	-0.287
Career Opportunities 2	0.2442	0.216	-0.2025	-0.4339	-0.2815
Compensation and Benefits 1	0.1717	0.5478	0.258	-0.0305	0.1613
Compensation and Benefits 2	0.1642	0.5336	0.3493	0.2523	0.2109
Knowledge Sharing 1	0.2354	0.0694	-0.0938	0.295	0.1228
Knowledge Sharing 2	0.2468	-0.0953	-0.1569	0.3326	0.1085
Employee Morale 1	0.2797	0.0701	-0.0523	0.1017	-0.1594
Employee Morale 2	0.2642	0.0866	-0.0458	0.1245	-0.1337
Feedback and Review 1	0.232	-0.1573	-0.0512	-0.2831	0.6022
Feedback and Review 2	0.2448	-0.167	-0.0329	-0.2823	0.5001
Sr. Leadership 1	0.2644	-0.1353	-0.2267	0.2709	-0.0337
Sr. Leadership 2	0.2593	-0.1272	-0.2179	0.3046	-0.0767
Work Life Balance 1	0.1993	-0.2748	0.566	-0.0377	-0.1693
Work Life Balance 2	0.2084	-0.3106	0.5278	-0.0457	-0.1733
Fairness and Respect 1	0.2559	-0.0818	-0.0226	-0.2024	-0.0451
Fairness and Respect 2	0.2682	-0.1744	0.0118	0.0322	-0.0544
Proportion Of Variance	0.5754	0.0655	0.0626	0.0449	0.0401
Number of obs	71,567				

Table II-7. Cross-Sectional Determinants of ESI

Following table shows the results of regression of Log ESI on firm characteristics. Dependent variable is log of the average of three-month employee satisfaction index (ESI) measured after the quarterly earnings announcement date. Firm characteristic variables are measured at the quarterly level at the time of the earnings announcement. Details of the accounting variables are shown in appendix II-2. Specification (1) is simple pooled regression. Specification (2) includes one digit SIC code dummies. Specification (3) is same as Specification (2) with firms in financial services industry excluded. Specification (4) is along the lines of Fama-Macbeth two stage procedure. Robust Standard errors are in parentheses. The sample period is from January 2008 to May 2011. ***, **, * denote 1%, 5% and 10% significance levels, respectively.

	Pooled Regression	**With Industry Dummies**	**Excluding Financial Services**	**Two-stage Fama-Macbeth**
Log ESI_t	**(1)**	**(2)**	**(3)**	**(4)**
Log $Assetst_{t-1}$	1.0921***	0.9423***	1.1157***	1.2526***
	(0.1145)	(0.1196)	(0.1274)	(0.2131)
Market to Book $Ratio_{t-1}$	0.8482***	0.8069***	0.9320***	1.2144***
	(0.2733)	(0.2691)	(0.2807)	(0.2030)
Leverage $(market)_{t-1}$	-3.8968***	-3.679***	-3.8141***	-4.2474***
	(1.1472)	(1.1725)	(1.2589)	(1.1809)
Return on $Assets_{t-1}$	-9.1466	-7.8971	-8.9222	8.0573
	(7.3481)	(7.3098)	(7.4314)	(10.5343)
SG&A over $Sales_{t-1}$	0.7664	0.3738	0.5477	2.5635
	(0.9424)	(0.8952)	(0.9208)	(1.8232)
Constant	41.142***	41.371***	40.875***	39.238***
	(1.1369)	(5.559)	(1.2546)	(2.3092)
Observations	6,586	6,586	6,083	7,752
R-squared	0.0275	0.03	0.0287	0.0422

Table II-8. Returns Around the Earnings Announcements

Specifications 1, 2 & 3 show the results of quarterly regression of the five-day cumulative abnormal return, centered on the earnings announcement date, on the lagged log Employee Satisfaction Index (ESI). ESI is the average of three-month employee satisfaction index (ESI) measured 80(-10,-80) days prior to the quarterly earnings announcement date (day 0). Abnormal return is calculated by subtracting the value weighted market return from the stock return. Accounting variables are based on the previous quarter's numbers: Log Assets is natural log of total assets. MarketToBook ratio is defined as the sum of total assets and market value of equity minus the book value of equity divided by total assets. Leverage is defined as long-term debt plus current liabilities divided by long-term debt plus current liabilities plus market value of common equity. ReturnOnAssets is income before extraordinary items divided by lagged assets value. SG&ARatio is the SG&A expense divided by lagged value of sales. Return60120 is the two-month return in the period from (-60,-120). Return3060 is the one-month return in the period from (-30,-60). Analyst Forecast Standard Deviation is the standard deviation of analyst's forecasts in the 90 day period prior to the earnings announcement.
Specification (4) excludes two quarters (last quarter of 2008 and first quarter of 2009) to account for financial crisis. Specifications 5, 6 show the results of the regressions run on small and large firms separately. The sample is divided into three bins based on total assets. "Small Firms" constitute the firms in the smallest size bin while the "Large Firms" constitute the firms in the largest size bin. Standard errors in parentheses are White standard errors clustered by firm. The sample period is from January 2008 to May 2011. ***, **, * denote 1%, 5% and 10% significance levels, respectively.

Five Day CAR					**Small Firms**	**Large Firms**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Log ESI$_{t-1}$	0.01369**	0.01378**	0.01446**	0.02026***	0.02564**	0.002715
	(0.0067)	(0.0067)	(0.0067)	(0.0075)	(0.0128)	(0.0107)
Log Assets$_{t-1}$	-0.02761*	-0.03195*	-0.03339**	-0.02880	-0.07200*	-0.05864**
	(0.0143)	(0.0163)	(0.0165)	(0.0184)	(0.0431)	(0.0263)
Market to Book Ratio$_{t-1}$	-0.01257**	-0.01346**	-0.01625**	-0.02031***	-0.02105	-0.03546***
	(0.0053)	(0.0058)	(0.0064)	(0.0077)	(0.0132)	(0.0092)
SG&A over Sales$_{t-1}$	0.01559**	0.01483**	0.01444**	0.009005*	0.01977**	-0.01970
	(0.0071)	(0.0066)	(0.0064)	(0.0050)	(0.0096)	(0.0270)
Return on Assets$_{t-1}$	-0.1040	-0.1134	-0.1119	-0.1703*	-0.05597	0.1352
	(0.1079)	(0.1057)	(0.1067)	(0.0921)	(0.1388)	(0.1413)
Analyst Forecast	-0.00001789	-0.00002079*	-0.00001862	0.000004367	-0.0008758	-0.00001913
Standard Deviation$_{t-1}$	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0057)	(0.0000)
Return3060			-0.02516	-0.005049	-0.01918	-0.007398
			(0.0160)	(0.0179)	(0.0299)	(0.0209)
Return60120			-0.02121**	-0.03745***	-0.04613**	-0.02845*
			(0.0107)	(0.0124)	(0.0219)	(0.0149)
Constant	0.1936	0.2315*	0.2905**	0.1829	0.4437*	0.6652**
	(0.1189)	(0.1336)	(0.1373)	(0.1564)	(0.2641)	(0.2613)
Observations	4,861	4,861	4,861	3,892	1,472	1,773
R-squared	0.0089	0.0152	0.0183	0.0235	0.0414	0.0266
Number of Firms	1,103	1,103	1,103	1,084	454	333
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	Yes	Yes	Yes	Yes	Yes

Table II-9. Human Capital Dependent Firms

Following table repeats the analysis of Table II-8 by dividing the sample according to proxies for human capital. In each of the specifications, dependent variable is the five-day cumulative abnormal return, centered on the earnings announcement date. The sample is divided into three bins based on each proxy for human capital. "Low" constitutes the firms in the smallest value bin while the "High" constitute the firms in the largest value bin. Specifications 1, 2 show the results of the regressions run on Low and High Market to Book ratio firms separately. Specifications 3, 4 show the results of the regressions run on Low and High R&D Ratio (R&D expense for the quarter divided by lagged value of total assets) firms separately. Specifications 5, 6 show the results of the regressions run on Low and High SG&A Ratio (SG&A expense for the quarter divided by lagged value of total sales) firms separately. Standard errors in parentheses are White standard errors clustered by firm. The sample period is from January 2008 to May 2011. ***, **, * denote 1%, 5% and 10% significance levels, respectively.

	Five Day CAR		Five Day CAR		Five Day CAR	
	Low Mkt. to Book Ratio	High Mkt. to Book Ratio	Low R&D Ratio	High R&D Ratio	Low SG&A Ratio	HighSG&A Ratio
	(1)	(2)	(3)	(4)	(5)	(6)
Log ESI$_{t-1}$	-0.009638	0.02486***	0.01475	0.03797**	0.01592	0.02593*
	(0.0161)	(0.0093)	(0.0148)	(0.0177)	(0.0115)	(0.0132)
Log Assets$_{t-1}$	-0.09487***	-0.04552	-0.03977	-0.08609**	-0.06071**	-0.02045
	(0.0329)	(0.0297)	(0.0362)	(0.0359)	(0.0303)	(0.0247)
Market to Book Ratio$_{t-1}$	-0.1671***	-0.01000*	-0.03032**	-0.005745	-0.03613***	-0.01405**
	(0.0523)	(0.0060)	(0.0124)	(0.0075)	(0.0107)	(0.0063)
SG&A over Sales$_{t-1}$	-0.002942	0.01395***	0.4936**	0.01416**		
	(0.0386)	(0.0043)	(0.1932)	(0.0056)		
Return on Assets$_{t-1}$	-0.1925	0.05215	-0.07499	-0.1463	-0.5319**	-0.004899
	(0.1721)	(0.1205)	(0.1669)	(0.1323)	(0.2245)	(0.1746)
Analyst Forecast Standard Deviation$_{t-1}$	-0.000009510	-0.000008949	-0.0004158	-0.00009795*	-0.000001034	-0.0001101**
	(0.0000)	(0.0000)	(0.0005)	(0.0001)	(0.0000)	(0.0001)
Return3060	-0.01575	-0.04774*	-0.02374	0.002202	-0.02241	-0.03851
	(0.0294)	(0.0245)	(0.0316)	(0.0448)	(0.0306)	(0.0316)
Return60120	-0.002937	-0.01758	-0.03416	-0.06521***	-0.01664	-0.01621
	(0.0199)	(0.0144)	(0.0233)	(0.0238)	(0.0189)	(0.0195)
Constant	1.0665***	0.3439	0.2758	0.5236*	0.5589**	0.1239
	(0.2884)	(0.2349)	(0.2896)	(0.2723)	(0.2701)	(0.1943)
Observations	1305	1919	1003	964	1577	1588
R-squared	0.0464	0.0238	0.0645	0.0592	0.0427	0.0312
Number of Firms	413	471	234	254	417	408
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Appendix II-1: Survey Questions

In a typical review an employee rates his/her company for the following are survey questions:

Survey Questions for IBM
How satisfied are you with...

Q1 Your opportunities for professional growth at IBM?

Q2 Your opportunities for career advancement within IBM?

Q3 Your compensation compared with similar jobs elsewhere?

Q4 Your benefits package compared with similar employers?

Q5 Information and knowledge sharing within IBM?

Q6 Communications from management about important issues and changes?

Q7 IBM as a place you would recommend to others to work?

Q8 IBM as a place you are proud to work?

Q9 Feedback you receive about your job performance?

Q10 Recognition and praise you receive when you do a good job?

Q11 Leadership abilities of Senior management?

Q12 Competence of Senior management?

Q13 Management support in permitting time off when you think it's necessary?

Q14 Employer support in balancing between work life and personal life?

Q15 Fairness in how promotions are given and people are treated?

Q16 The level of respect shown by management toward employees?

Q17 Overall, how satisfied are you with IBM as a place to work?

Employees rate each of the questions on the following scale:

1 - Very Dissatisfied 2 - Dissatisfied
3 - Neutral
4 - Satisfied 5 - Very Satisfied

Appendix II-2: Variable Definitions

Main accounting variables used in the paper are defined as follows:

Log Assets is natural log of total assets (Compustat item actq).

MktValue is market value of common equity and is defined as common shares outstanding times the quarter end price (Compustat cshoq*prccq).

ReturnonAssets is income before extraordinary items divided by lagged assets value (Compustat ibq/l.atq)

Market to Book Ratio is market to book ratio and is defined as the sum of total assets and market value of equity minus the book value of equity divided by total assets (Compustat (atq-seqq+cshoq*prccq)/atq).

LeverageMkt is defined as long-term debt plus current liabilities divided by long-term debt plus current liabilities plus market value of common equity (Compustat (dlcq + dlttq) / (dlcq + dlttq + cshoq*prccq)).

SG&A over Sales is SG&A expense for the quarter divided by lagged value of sales (Compustat xsgaq/sale).

R&D over Assets is R&D expense for the quarter divided by lagged value of total assets (Compustat xrdq/l.atq).

Analyst Forecast Standard Deviation is the standard deviation of analyst's forecasts in the 90 day period prior to the earnings announcement.

Return3060 is the 30 day return from day -30 to day -60. Day 0 is the earnings announcement date

Return60120 is the 60 day return from day -60 to day -120. Day 0 is the earnings announcement date

References

Abowd, J., 1989. The effect of wage bargains on the stock market value of the firm. *American Economic Review* 79, 774–800.

Akerlof, G., 1982. Labor contracts as partial gift exchange. *Quarterly Journal of Economic*s 97, 543–569.

Akerlof, G., Yellen, J., 1986. Efficiency Wage Models of the Labor Market. Cambridge University Press, Cambridge, UK.

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, 785-818.

Battalio, R.H., and R. R. Mendenhall, 2005. Earnings expectations, investor trade size, and anomalous returns around earnings announcements, *Journal of Financial Economics* 77: 289-319.

Doyle, J.T., R.J. Lundholm, and M.T. Soliman. 2005. The Extreme Future Stock Returns Following Extreme Earnings Surprises. Working Paper, Stanford University.

D. W. Diamond and R. E. Verrecchia. Information Aggregation in a Noisy Rational Expectations Economy. *Journal of Financial Economics* 9 (September 1981), 221-35.

Edmans, Alex, 2009, Does the stock market fully value intangibles? Employee satisfaction and equity prices. *Journal of Financial Economics* 101 (2011) 621–640

Faleye, O., Trahan, E., 2006. Is what's best for employees best for shareholders? Unpublished working paper, Northeastern University.

Filbeck, G., Preece, D., 2003. Fortune's best 100 companies to work for in America: Do they work for shareholders? *Journal of Business, Finance and Accounting* 30, 771–797.

Foster, G., C. Olsen, and T. Shevlin. 1984. Earnings Releases, Anomalies, and the Behavior of Security Returns, *The Accounting Review*: 574-603.

Francis, J., K. Schipper, and L. Vincent. 2002. Expanded Disclosures and the Increased Usefulness of Earnings Announcements, *The Accounting Review* 77: 515-546.

Gorton, G., Schmid, F., 2004. Capital, labor, and the firm: a study of German codetermination. *Journal of the European Economic Association* 2, 863–905.

Imhoff, E.A. 1992. The Relation between Perceived Accounting Quality and Economic Characteristics of the Firm, *Journal of Accounting and Public Policy* 11: 97-118.

Lev, B. 1989. On the Usefulness of Earnings and Earnings Research: Lessons and Directions from Two Decades of Empirical Research, *Journal of Accounting Research* 27: 153-192

Liu, J., and Thomas, J. 2000. Stock Returns and Accounting Earnings, *Journal of Accounting Research* 38: 71-101

Livnat, J., and R.R. Mendenhall. 2006. Comparing the Post-Earnings Announcement Drift forSurprises Calculated from Analyst and Time-Series Forecasts, *Journal of Accounting*

Research 44: 177-204.

McGregor, D., 1960. The Human Side of Enterprise. McGraw-Hill, New York

Chapter III

Private Equity Ownership and the Performance of Reverse Leveraged Buyouts

Abstract

Using a hand collected sample of 133 reverse leveraged buyout firms from 1997-2002, I examine the financial performance of the firms immediately before the IPO and up to four years after the IPO. I find that for three years after the IPO they continue to outperform their industries. However, performance deteriorates after the IPO. Cross-sectional regression at time of the IPO suggests that long term performance after the IPO is related to changes in ownership by the private equity sponsors and is not related to changes in ownership by other insiders (all officers or directors who are not PE sponsors) or change in leverage. Even after the IPO, I find the positive relation between PE sponsor ownership and future performance continues. To establish causality between PE sponsor ownership and future performance I use the 2SLS-IV approach. The identifying instrument is the number of years since LBO and is a proxy for impatience of PE sponsors to free up their capital. I find that an IV regression finds a weakly significant relation between PE sponsor ownership and future performance.

3.1 Introduction

In recent years private equity backed deals have become an important part of US public markets. According to Morgan Stanley private equity backed leveraged buyout (LBO) volume has surged from $12bn in 1994 to $458bn in 2006. In the same year, 20.2% of M&A market was composed of private equity.

Private equity firms such as KKR, Carlyle Group or Thomas H. Lee Partners frequently participate in the public markets by conducting leverage buyouts (LBO) of the companies. The target of leveraged buyouts can be mature and out of favor public companies or their divisions, or even privately held and family-controlled companies. Private equity sponsors (PE sponsors)[5] maintain control of the firm for few years and then, at a later date, some of these LBO firms are taken to public markets through an IPO. This process of taking a company private only to return it to public markets again at a later date is called reverse leveraged buyout (RLBO).

I look at the financial performance of a hand- collected sample of 133 RLBO firms up to two years before the IPO and up to five years after the IPO. I find that these firms outperform their industry averages for up to three years after the IPO. Performance as measured by ROA peaks one year after the IPO and then declines thereafter. Tobin's Q is highest in the year of the IPO and also declines thereafter. These findings are broadly consistent with the findings of Holthausen and Larcker (1996) who, using a sample of 90 RLBOs between 1983 and 1988, also find a decline in operating performance of the companies after the RLBO.

[5] I use the terms private equity firm or private equity sponsors (PE sponsor) interchangeably in this paper to refer to the private equity firm that was instrumental in conducting the LBO.

Holthausen and Larcker find that the decline in long term performance of RLBOs is related to the decrease in ownership of management and non-management insiders. However, they treat all non-management insiders as homogenous. PE sponsors are considered non- management insiders in their paper, but they also include in this category any other board member or officer who is not part of the operating management. Therefore, their group of non-management insiders can include officers or directors representing diverse groups, such as institutional investors, family members, representatives of banks, representatives of various trusts and other insiders such as ex-CEO.

In this paper, I focus only on PE sponsors and their incentives to monitor the firm. I contend that PE sponsors and the other non-management insiders can have differences in their motivations and objectives in the firm. For example, PE sponsors and representatives of banks represent different classes of stakeholders' altogether. Family members and other directors such as the ex-CEO, who were executives of the firm earlier, will behave more similarly with current management than with PE sponsors. PE sponsors would like to exit as soon as possible from the firm after an IPO to realize their gains and free up the capital in order to return it to the limited partners or recycle it for future LBOs. But, managers and other board of directors may well be more motivated by entrenchment reasons. It should also be noted that other insiders group, excluding PE sponsors, holds substantial stake in the firm before and after the RLBO.

PE sponsors are typically seasoned industry veterans who have built their reputation by turning companies around. Having such experienced veterans as owners and board members of the company can have a beneficial effect in terms of better monitoring of the management and advise they impart to the management. According to the article "Perform or Perish" (Business Week, November5, 2007):

> "But along with the upside potential [refers to cash-out at IPO by executives at the LBO firm] comes constant accountability to private equity bosses, near-impossible

> time-tables, ever increasing financial goals and grueling board meetings in which standout numbers are met with blank faces…."We expect CEOs to exceed our plan, and we talk to them daily or every other day," says one chairman and CEO of a private equity firm."

From the above statement it would seem that PE sponsors monitor and advise management diligently and this may lead to value creation in the company.

However, several articles in the financial press and various lawsuits contend that PE firms harm the firms by overleveraging them, by paying themselves special dividends at expense of other stakeholders, by cutting jobs, and by decreasing the long term investments in capital expenditure and R&D. As PE sponsors take firms public and stand to cash-out, their interests are not aligned with those of the other shareholders. Chou, Gombola, and Liu (2006) analyze earnings management around RLBOs and find positive and significant discretionary current accruals coincident with RLBOs. However, reputational concerns limit the extent of opportunism by PE sponsors.

Using ownership of PE sponsors as proxy for their incentives, I test the hypothesis that, higher the ownership stake is, more likely are PE sponsors to provide beneficial monitoring and advice to the management. PE sponsors reduce their ownership at the time of IPO by 29% on average with respect to their ownership immediately prior to IPO. I find that cross-sectional variation in long term firm performance after the IPO is positively related to changes in ownership by PE sponsors. It can also be argued that PE sponsors reduce their stake by a greater amount when they know that the company is going to perform badly in future. To test the above hypothesis, I look at the changes in ownership of other insiders (all officers or directors who are not PE sponsors) and find that they reduce their ownership by about 23% on average with respect to their ownership before the IPO. However, while the magnitude of percentage change in ownership of the other insiders and PE sponsors is comparable, yet there is no significant relation between changes in other

insiders' ownership and changes in performance. It is hard to reconcile that other insiders are not decreasing their stake for information asymmetry reasons while PE sponsors are. I also argue that, because PE sponsors are repeat players in the public markets, it is in their long term interest to not trade due to information asymmetry reasons.

I cannot interpret these results to imply that decrease in ownership of PE sponsors is causing the decrease in long term performance, as this implicitly assumes that changes in ownership are exogenous. It is possible that some exogenous change in the economic environment that affects both ownership and performance is causing the observed association between ownership and performance. Hence, the changes in ownership and performance can be endogenous.

I try to ameliorate the endogeneity problem between ownership and performance using a panel data specification with firm fixed effects and two stage least squares- instrument variable approach. I construct a panel data set where I have ownership and performance information for the RLBO companies for up to seven years after the IPO. As pointed out by Himmelberg, Hubbard, & Palia (1999), if one assumes that unobserved heterogeneity is a firm fixed effect, one can use firm fixed effects specification to reduce endogeneity concerns. Secondly, I use lagged values of independent variables to overcome any contemporaneous variables affecting both ownership and performance.

Finally, I use the total amount of time elapsed since RLBO as a proxy for the impatience of PE sponsors. PE firms are contractually bound to return the capital of limited partners within 5 to 10 years (Sudarsanam 2003, Ch. 12). Failure to do so in a timely manner can have a negative effect on the PE firm's reputation. Bad reputation can make it difficult for a PE firm to raise capital in the future. Hence, PE sponsors have to sell their stakes in the companies post-IPO to generate sufficient funds to be able to return the limited partners' capital. Moreover, sooner the PE firms are able to free the capital from one RLBO firm, the sooner they can use that capital to invest in other

LBO deals. Using total time since LBO as an instrument, I find that as PE firms decrease their ownership the future operating performance also declines. It seems that as ownership of PE sponsors declines they have less incentive to monitor and advise the management and this is related to poor performance in the future.

The remaining part of the paper is organized as follows. Section 2 gives a brief summary of existing literature followed by a description of the data and considers potential sources of bias. Section 3 looks at performance of firms before and after the RLBO. Section 4 presents results and analysis of key findings. Section 5 summarizes and concludes.

3.2 Review of Literature and Sample Description

3.2.1 Literature Review

Several papers in the earlier literature, mostly drawing on data from the 1980s have explored the performance of RLBOs. Degeorge and Zeckhauser (1993) examined 62 RLBOs between 1983 and 1987, and find that the accounting performance of these firms exceeds the average industry performance prior to going public, and then deteriorates afterwards.

Holthausen and Larcker (1996) also find weak evidence that the accounting performance of RLBOs decreases after they go public. They look at the effect of change in management and other non-insiders' ownership immediately before and after the RLBO of 90 companies from 1983 to 1988 and its effect on accounting performance post RLBO. However, as I do in this paper, they do not specifically look at PE sponsors ownership and the changes in the PE sponsors' ownership after the RLBO. Since PE sponsors are considered the main catalyst in transforming the RLBO companies, it is natural to look at only the ownership of the PE sponsors without any contamination

by the other board members ownership. Mian and Rosenfeld (1993) examine 85 RLBOs over roughly the same period and find that the offerings slightly outperform the benchmark portfolio.

More recently, Cao and Lerner in their working paper find that, in terms of risk adjusted stock returns, RLBOs consistently outperform other IPOs and stock market as a whole over three to five year periods. Cao has a working paper in which he looks at the performance of RLBOs and the exit strategies of RLBO sponsors. His findings also show that RLBOs have significantly lower under-pricing than other IPOs. Similar to these papers I look at post-IPO performance of RLBOs, however I focus on changes in ownership of PE sponsors and its effect on future performance of the RLBO firm at time of the IPO and up to seven years after the IPO.

There are two well cited studies about post-IPO performance and managerial ownership. Jain and Kini, 1994 find positive link between management ownership and post-IPO performance. However, Mikkelson, Partch and Shah, 1997 looking at post-IPO performance for up to 10 years find no link between ownership and operating performance of companies that go public. Different from these papers I specifically look at effect of private equity ownership on performance of RLBOs, while controlling for managerial ownership.

3.2.2 Data

There has not been much research in the area of RLBO in last 10 years. Research in the 1990s, for most part, drew upon a sample of 90 RLBOs provided by Kidder, Peabody & Co. Due to various considerations such as unavailability of data related to performance or governance, most of the analysis drew upon a sample of 40 or 50 of these 90 RLBOs. Hence, one of the important hurdles that a researcher has to overcome in this field is identifying the RLBO sample.

I used various databases such as SDC Platinum, Mergerstat, CapIQ, Zephyr (Bureau Van Dijk) to identify companies in the US that underwent a leveraged buyout with help of a private

equity sponsor. I excluded deals that involved no PE sponsor or were characterized as a management buyout or were venture capital related. This gave me a list of about 600 companies. I then checked Hoover's Company Records and Mergent Online for company history for each of these companies to see if these companies had undergone an IPO at a later date and before end of the year 2002. The choice of the year 2002 was made to ensure that I have at least four years of post RLBO performance and ownership data.

I also checked against the new issues database in SDC Platinum to see if the company subsequently conducted an IPO. However, this resulted in very few hits as many of the RLBO companies change their names/ticker when they undergo an IPO at a later date. In the end, I had an LBO–subsequent IPO matched sample of about 98 companies. My sample is different from the sample used in the papers in 1990s as the sample of companies used in earlier studies underwent RLBO before 1989, while my list of companies underwent RLBO from 1997 to 2002.

3.2.2.1 Sample Selection Bias Issues

It is possible that many of the databases are more likely to have coverage of larger and relatively well known companies; this will bias my sample towards larger companies. To further expand list of my companies and to ensure that I was not systematically missing any RLBOs, I compared my list with the list of RLBOs, with offering size greater than $50 million, kindly provided to me by Josh Lerner and Jerry Cao. In their NBER working paper, The Performance of Reverse Leveraged Buyouts, they analyze the stock performance of 496 RLBOs between 1980 and 2002. Their sample dataset is very comprehensive and to eliminate any systematic biases they even ask PE groups to check their sample and to confirm that it is representative of the companies they backed in RLBO.

Thus, adding another 48 companies from the Lerner and Cao list of RLBOs, I finally have a list of 133 companies that conducted RLBO between 1997 and 2002, were backed by private equity group and had offering size of at least $50 million at time of the RLBO. For purposes of calculation

of performance of RLBOs, I excluded companies that were utilities or financials and this left me a sample of 116 companies. Table III-1 shows the distribution of IPO completions dates and their LBO dates for the sample of 133 RLBOs between 1997 and 2002. Number of RLBOs steadily increased from 1997 to 1999, peaked in 1999 and 2000 and subsequently declined till 2002. This indicates that RLBOs tend to follow economic cycles.

Table III-2 shows distribution of sample RLBO firms grouped by SIC codes. While manufacturing as an industry had the highest number of RLBOs, it does not appear that RLBOs are concentrated only in select industries. This finding is consistent with earlier literature drawing on sample of RLBOs completed before 1989. Table III-3 presents the summary statistics for the sample of RLBO companies.

Data on performance was taken from Compustat and is by fiscal year. Data for performance and ownership before the IPO was taken from the latest amendment to S-1 registration statements, filed with SEC, prior to the IPO. Data on ownership after the IPO was collected by looking at proxy statements up to 8 years after the IPO. In calculation of percent ownership, I include shares of common stock that can be issued upon the exercise of outstanding options and warrants that are currently exercisable or exercisable within 60 days of the date of proxy statement.

Period of 1997 and 2002 was marred by an internet bubble and a subsequent recession; it is possible that my results are being affected by this volatile time period. The impact of this time period is lessened in my study because of several reasons; RLBOs are typically of companies that are mature and less prone to "irrational exuberance". Moreover, in my sample I do not have any internet firms and there are only 12 firms that are in someway related to technology sector. Finally, to overcome any industry wide fluctuations I only look at industry adjusted performance measures. Specifically, I control for industry and year effects.

3.3. Performance of RLBOs

3.3.1 Long Term Performance of RLBOs

I look at three different performance measures to assess performance of the firms after the RLBO. Tobin's Q is a performance measure that is forward-looking and is most widely used in the papers analyzing the link between ownership and performance. The other two performance measures, I use, are the return on assets and return on sales. Appendix III-1 details the construction of these measures. As was done in earlier literature, I use industry averages as benchmark for performance. Specifically, the benchmark is the mean adjusted industry measure for each fiscal year. Industries are grouped according to the three digit SIC code. Performance measures are skewed towards larger companies and hence median adjusted benchmark may be more appropriate. However, as I limit my dataset to relatively large companies (equity offering at least $50 million) it may be more appropriate if I use mean adjusted measure that intrinsically compensates for the size. I also ran my results with median adjusted numbers and get similar results. Some firms change their SIC codes over the years. In such cases, I use historical SIC code (data item 277) in Compustat to ensure that I classify the firm in correct 3 digit SIC code for each year.

Return on assets is defined as earnings before depreciation and interest divided by the book value of assets. Using earnings before interest also overcomes the mechanical effect of leverage on the income. One of the potential problems with deflation by assets, as pointed by Holthausen and Larcker, is that LBO firms are likely to have written up the assets or created goodwill, in accordance with GAAP, to the extent that the buyout price at time of IPO exceeds the book value of assets at the time of IPO. If the write up of assets and goodwill by an LBO firms exceeds that of the industry in general, the performance of LBO firms will be downward biased. As an alternative, I also provide the performance deflated by sales in the table. However, deflating by sales may not capture the

efficiencies that management created by increasing sales without proportionately increasing the assets. As both measures suffer from some drawbacks, I report both these measures for comparison. The general conclusions are similar for both measures and are reported in the Table III-4.

Each variable is measured for each firm for up to two years before the IPO and up to five years after the IPO. Year zero is the calendar year of the IPO. I do not control for the month of the fiscal year. For some firms, the performance in year zero may largely be driven by the time when firm was private while in others it may be driven by the time it was public. The number of firms for each year also varies. Some of the firms go bankrupt or are taken over after the IPO. These firms do not appear in years subsequent to their takeover or bankruptcies. For some firms, I was not able to get performance data from their S1 registration statements or Compustat; these firms also do not appear in these numbers.

In general, the industry adjusted performance measures indicate that RLBO firms perform significantly better than their industry averages for up to three years after the IPO. For example, the median industry adjusted EBITDA/Assets is 0.048 in the year of IPO and 0.072 in the IPO+1 year. Thus a median RLBO firm returned .048 more cents per dollar of assets as compared to an average firm in its three digit SIC industry. Similarly, a median RLBO firm returned .072 more cents per dollar of assets in the year after the IPO as compared to the industry average. These numbers are consistent with earlier studies in the literature that document performance of RLBO firms.

I cannot interpret these results to say that LBO firms perform better than other similar firms. There is selection bias inherent in any studies of the LBOs since not all the LBOs eventually return to the public ownership. It can be argued that weakly performing LBOs are not taken public. Therefore, the sample of RLBOs or the LBOs that are taken public will be biased towards better performing firms.

Peak operating performance is reached in the year after IPO year. In earlier studies by Holthausen and Larcker and Degeorge and Zeckhauser, based on RLBOs between 1977 and 1989, it was documented that peak operating performance was reached in the year of the IPO. This was interpreted as support for market timing hypothesis and managers taking advantage of information asymmetries to take firm public when they know that peak performance is reached. My results, however, indicate that performance improves for one full year after the IPO. It is possible that as many of the RLBOs are subsequently taken over, it is in interest of the PE sponsors to ensure that firms continue to perform well after the IPO. This view is consistent with Mian and Rosenfeld (1993) and Muscarella and Vetsuypens (1990) who find that PE sponsors use IPO as opportunity to show-case the firm, and this that may aid in its subsequent sale to an outside buyer.

Another possibility is that since in late 1980s and early 1990s many RLBOs performed poorly after the IPO, which in turn led to several law suits against the management and PE sponsors of these failing companies. The threat of lawsuits can possibly lead risk averse PE sponsors to take the firms public when they have a good chance of showing improved performance year after the IPO.

I also find support for the view that PE sponsors successfully time the IPO market. Tobin's Q is at its highest level in the year of the IPO, which can imply that the companies successfully timed the market and undertook IPO in the year when the market valued them the most, even though operating performance had not peaked yet. Moreover, as the market was in a bubble phase in part of the period from 1997 to 2002. This may have lead PE sponsors to take the companies public even though all the planned improvements in the company had not been made, as delaying the IPO could had meant foregoing a favorable window of opportunity.

Hence, it seems that the motivation to take companies public was driven more by the market dynamics rather than by the peaking operating performance. Holthausen and Larcker find, in their

sample of RLBOs, that the firms were taken public in the year when operating performance had peaked. However, they do not report Tobin's Q for their set of RLBOs[6].

Table III-4 also shows that there is deterioration in performance after the IPO and operating performance steadily decreases after the IPO+1 year. Performance as measured by Tobin's Q peaks at the year of IPO and then continues to decline over the next four years. In general, I find that while RLBOs continue to outperform their industries for up to three years after the IPO, they do show a decline in performance after the IPO.

3.3.2 Reasons for decline in performance:

The decline in performance can be due to several different reasons. Major value drivers of the LBOs are the disciplining effect of higher leverage, monitoring by PE sponsors and incentive alignment of management through greater shareholdings. After a RLBO all of the above value drivers decrease substantially. Based on theoretical considerations I summarize the effect of each of these value drivers on performance of the firm below.

3.3.2.1 PE Sponsor Ownership

Table III-5 details the ownership of PE Sponsors and management (excluding the PE sponsors) before and after the RLBO. I was able to obtain ownership data for only 73 of 116 RLBO companies from the proxy statements and IPO prospectus.

PE sponsors on average decrease their ownership by 16.9% in absolute terms, with the maximum being 84.1%. Even after RLBO they continue to have significant board presence and several of them are chairman of their respective boards and are also part of key committees such as the compensation committee. If primary role of a PE

[6] I have unsuccessfully tried to get the list of 90 RLBOs used in earlier studies from the various authors. Since this dataset is almost twenty years old, none of the authors I contacted seem to have it.

sponsor is to monitor and advise the management, it can be treated as a fixed cost and, hence, a reduction in ownership leads to less benefit from any value maximizing actions that they may take. This will lower their incentives to incur the fixed cost of monitoring and advising the management. Thus a reduction in their ownership can have a negative effect on the firm performance.

It is also possible that PE sponsors indulge in short-termism at the expense of long term value of the company. However, there are reputational concerns that mitigate these problems. It is in interest of PE sponsors to be seen as creators of shareholder

wealth, as this will allow them to conduct future LBOs and their subsequent IPOs more easily. The ability to easily conduct future LBOs is even more important now-a-days as most of their target companies are likely to have anti-takeover defenses that preclude the hostile takeovers of 1980s. Moreover, as can be seen from the Table III-5 the PE sponsors continue to hold on average 28.1%, a substantial percentage, of company even three years after the IPO. Thus any short term behavior that has negative impact on the long term value of the company will also hurt them as the value of their holdings will also decline.

3.3.2.2 Management or Insider Ownership Excluding PE Sponsors

After RLBO, the other insiders (composed primarily of management of the company) reduce their stake in the company by 6.9% percent on average in absolute terms. The percentage decrease is 23% with respect to their holdings immediately prior to IPO. Higher ownership by managers helps to align the incentives of management with the owners of the company and can lead to better firm performance (Jensen and Meckling (1976)) as managers get a higher portion of value creation in the company. However, it is also possible that higher ownership by managers leads to lower financial performance since risk averse, under-diversified managers avoid positive NPV high risk projects, Fama (1980) and Amihud and Lev (1981). Managerial risk aversion and under-diversification coupled

with high leverage can lead to even more risk averse behavior by managers as high leverage increases the probability of bankruptcy and, hence, the probability that managers may lose their jobs.

3.3.2.3 Leverage

Median industry adjusted leverage for RLBOs reduces from 35% above the mean industry levels to only about 4% above the mean industry levels after the RLBO. Excessive leverage motivates management to improve operating performance and use the free cash flow to make high interest payments which reduces the agency costs related to free cash flow (Jensen 1986). However, excessive leverage can also lead to negative incentive affects as management may take sub-optimal actions such as asset substitution. Managers acting in interest of shareholders may chose to take high risk projects that benefit shareholders at expense of bondholders (Jensen Meckling (1976)). Moreover, higher leverage may induce risk averse managers to invest only in low risk projects to avoid possible financial distress. Novaes and Zingales (1995) show that the managers motivated by entrenchment reasons will chose a different debt level from the one that is optimal for the shareholders.

From above discussion, we can see that the impact on performance of decrease in leverage and ownership stakes by management and PE sponsors is not clear and is an empirical question that I address in the next section.

3.4. Empirical Analysis

3.4.1 Cross-Sectional regression of performance on ownership

To empirically test the impact of ownership and leverage changes on long term performance, I average Tobin's Q for up to two years after the IPO and proxy it for long term performance. Since

it is not clear when the future performance is affected by the change in ownership, I average the future performance over two, three and four years after the IPO.

Change in PE sponsor ownership, immediately before and after the IPO, is used to proxy for the incentives that PE sponsors have to monitor the management. The other independent variable is the change in leverage from the year before IPO to the year after IPO. To check the effect on performance by change in incentives of management, I also examine the change in ownership by management. Ownership of management variable includes holdings of all executives and board of directors excluding the PE sponsors.

The cross-sectional regression of performance on ownership at time of IPO is specified as follows:

$$AvgPerf_i = a + \beta_1 ChngPEOwn_{i,(IPO-1\ to\ IPO+1)} + \beta_2 ChngMgmtOwn_{i,(IPO-1\ to\ IPO+1)} + \beta_3 ChngLev_{i,(IPO-1\ to\ IPO+1)} + ControlVariables + \varepsilon_i \text{ , where } i=1,2,\ldots..,N.$$

If the long run performance is adversely affected by reduction in ownership and reduced leverage, one would expect significant positive coefficients for β_1, β_2 and β_3 .

As shown in Table III-6, only the reduction in PE sponsors ownership is significantly positively related to the decrease in performance. To understand the economic significance of the results I examinee the effect of one standard deviation change in PE sponsors ownership on performance. A one standard deviation decrease in PE sponsors ownership is a decrease by 12.7% in absolute terms. This reduction is related to decrease in performance by 12.7*.0166, or about 0.21. The mean of the industry adjusted Tobin's Q averaged over two subsequent years is 0.499. Hence, one standard deviation reduction in PE ownership is positively related to about 42% reduction in average performance. It appears that, as ownership stake of PE sponsors decreases, they have less incentive to monitor and this is related to deterioration in the performance of these RLBO firms after the IPO. Given that most of the LBO candidates suffer from agency problems and free cash

flow problems, it is plausible, that as the amount of monitoring is reduced, there is tendency to revert back to pre-LBO problems. Since leverage and ownership of other insiders can be endogenous, I do not interpret the results to imply that there is no relation between other insiders' ownership and performance or leverage and performance.

3.4.2 Alternative Explanations

I consider four alternative hypotheses that are consistent with the observed positive relation between the ownership of PE sponsors and the future performance of the firm. The first two explanations rely on information asymmetry arguments while the third one brings forth the possible endogeneity between ownership and performance. Final alternative explanation considers possible earnings management at the time of the IPO.

It is possible that in presence of information asymmetry, PE sponsors hold higher stakes in good companies to signal the quality of the RLBO firms at the time of the IPO (Leland and Pyle (1977)). Second, as PE sponsors are considered insiders, it is possible that they sell off their stakes when they have inside information about some bad news affecting the firm performance. Finally, it is also possible that there are some observed and unobserved variables in a firm's contracting environment that are related to both ownership and performance (Himmelberg, Hubbard, & Palia (1999)). In such cases also one can get spurious positive relation between ownership and performance.

Leland and Pyle (1977) contend that entrepreneurs hold substantial stake in companies to signal the quality of the companies. PE sponsors may hold more stake in companies that they expect to perform well in future and, thus, use their ownership stake as a signal to public markets. In such cases, one can get a positive relation between ownership and performance. In general, it is difficult to differentiate between the monitoring and signaling hypothesis, because the predictions of both

these hypothesis are the same. However, there is some evidence that my results are probably more driven by the monitoring hypothesis. First, Leland and Pyle predict that management should hold more in companies that are expected to have good performance. But, in my sample other insiders that include the management of the company do not hold more in good companies. This is all the more significant because the percent reduction in insider holdings posts IPO is 23% and is comparable to percent reduction of 29% by PE sponsors. Leland and Pyle's results are based on the information asymmetry argument and, since RLBOs were public at an earlier date and on average tend to be larger companies, it can be argued that they suffer less from information asymmetry problems as compared to other IPOs, (Bharat and Kini, 1994). Finally, if there is no beneficial effect of PE Sponsor ownership and my results are being driven exclusively by signaling hypothesis, then I should expect no relation between performance and ownership post-IPO. In the panel data specification below I find that, even post-IPO, the positive relation between ownership and future performance continues.

An alternative way to interpret the results will be to consider that PE sponsors are insiders and they sell off more of their stake in the companies that they know are going to perform badly in future. This hypothesis differs from the signaling hypothesis as it assumes that, somehow PE sponsors are able to fool the market and indulge in insiders trading. Thus, I would expect that other insiders, who include executives and other directors, will also try to take advantage of the insider information and will tend to sell of their shares accordingly. However, in the above regression, I find that the selling by other insiders is not related to the future long term performance. Hence, both these groups of insiders are substantially reducing their holdings post IPO and it seems implausible that one group of insiders would take advantage of information asymmetry while the other group would not.

Secondly, such opportunism can cause the stock price of the company to suffer as market interprets these signals. PE sponsors hold on average 38% of the firm even after the IPO and it is not in their interest to trade due to information asymmetry reasons as this will cause them to lose substantial amount in value of their shares that they continue to hold.

Finally, a majority of PE sponsors are repeat players in the public markets. Typically they will have few other companies that they are getting ready for IPO and in future also they will be conducting LBOs and their subsequent IPOs. If the market discerns that PE sponsors trade on basis of insider information and some bad news is released soon after they sell, then in all the future deals PE sponsors stand to lose whenever they sell their shares in the market. PE sponsors cannot hold on to the shares of these companies indefinitely since they have to return capital to limited partners within a specific time frame. Therefore, as PE sponsors know that they will conduct future IPOs and then sell shares subsequently, it is in their interest to not develop reputation for trading due to insider information. While I find some evidence against the above mentioned information asymmetry hypotheses, still, it is difficult if not impossible to disentangle these alternative hypotheses.

It is also possible that earnings management may be influencing the results. The pattern of peak performance around IPO and declines thereafter may be caused by earnings management at time of the IPO. At time of the IPO, the information asymmetry problems are most severe and hence earnings manipulation should be most prominent at that time. Specifically, I look at firms that are in the top quartile by industry adjusted ROA in the year of the IPO. I assume that firms that have ROA in the year of IPO much higher than the industry average are more susceptible to charge of earnings management, than firms that have lower ROA. If earnings manipulation is driving the results then I would expect that for this sample of top ROA firms the results should be stronger than the rest of the sample. However, I find that results are weaker for this sample of firms (Table

III-7). Moreover, the behavior of other insiders in this subsample is even more dramatically opposed to the PE sponsors. If management had indulged in earnings management, one should expect them to sell to take advantage of it. As Table III-7 shows there is no relation between selling by other insiders, that include the CEO and other top executive officers of the firm, and the performance of the company. This indicates that earnings management may not be driving these results. Cao and Lerner in their 2007 working paper find their sample of RLBOs (my sample has similar companies from 1997-2002) consistently outperform other IPOs and stock market as a whole over three to five year periods. They break down their sample by year and hence I am able to confirm that the RLBOs outperformed the market on risk adjusted basis in years 1997 to 2002. Similar results for RLBOs are documented by Holthausen and Larcker (1996). If markets are semi strong efficient we do not expect that the firms that indulge in earnings management to outperform the market. Moreover, reputation concerns should limit this kind of behavior as PE groups are repeat players in the IPO markets.

Moreover, as shown in next section, the effect of ownership on future performance continues for up to seven years after the IPO. Once the firm is public information asymmetries are greatly reduced and alternative explanations to monitoring hypothesis predict no relation between ownership and future performance.

3.4.3 Endogeneity Concerns

As pointed by Himmelberg, Hubbard, & Palia (1999), it is possible that there are some observed and unobserved variables in a firm's contracting environment that are related to both ownership and performance. In such case my results can be driven by the variables absent in my specification and the link between PE sponsor ownership and performance will be spurious. Himmelberg, Hubbard, & Palia contend that using a panel data specification can ameliorate the endogeneity concerns if we

assume that unobserved heterogeneity is a firm fixed effect. Appendix III-2 details three examples of unobserved heterogeneity that can induce positive spurious correlation between ownership and performance.

I collected ownership data from the proxy statements for each firm for up to seven years after the IPO or till the time when PE sponsors had completely exited from the company. Thus, I have an unbalanced panel data where for each firm I have the ownership data as well as the performance data for up to seven years after the IPO. As PE sponsors continue to substantially decrease their stakes over the next several years, I would expect the positive relationship between ownership and performance to continue over this extended time period.

To ameliorate the endogeneity problem between ownership and performance I use following specification. If one assumes that unobserved heterogeneity is a firm fixed effect, one can use firm fixed effects specification to reduce endogeneity concerns. Secondly, to control for observed and unobserved variables that are contemporaneous and affect both ownership and performance, I lag the independent variables by one period, and consider their effect on performance in the next two periods. Dependent variables are the Tobin's Q and ROA, and are averages of the two subsequent periods. I also ran the regression using average performance over next three periods as dependent variable, and got similar results, however, this significantly reduces the number of observations and hence reduces the power of the tests. The specification is as follows:

$$AvgPerf_{i,t+1,t+2} = a_i + \beta_1 PEOwn_{i,t} + \beta_2 MgmtOwn_{i,t} + \beta_3 Lev_{i,t} + \beta_4 Logsales_{i,t+2} + \varepsilon_{i,t+1,t+2} \quad \text{,where } t = IPO+0, IPO+1,\ldots\ldots,IPO+6 \text{ and } i=1,2,\ldots..,N.$$

Results of the above specification are presented in Table III-8. The only coefficient that is significant is the PE sponsors ownership. Hence, even in presence of unobserved heterogeneity that is constant for a firm or is contemporaneous, I find PE sponsor ownership is positively related to performance.

It is still possible that there is some other variable that is affecting performance and PE sponsors ownership. Using the two stage least squares instrument variable approach, I try to overcome this concern. I use the fact that PE sponsors would like to free their capital as soon as possible post-IPO and use it to repay their limited partners. I use the variable, number of years since LBO, as proxy for their impatience to free up their capital. I am implicitly assuming that PE sponsors are reducing their ownership in the firms driven by their contract with limited partners that stipulates that they return their capital within a specific time frame, usually 5-10 years or due to emergence of opportunities to conduct some other deals[7]. The above motivations, while related with their ownership declines, have no specific relation with performance.

Ex-ante there is no clear reason to expect any relationship between performance and time since LBO. The life cycle argument used to explain performance of IPO's may not hold for RLBOs. RLBOs tend to be of companies that were once public and are in low growth mature phase of their lifecycles. Hence, the life cycle variation should be negligible in these mature firms. Jain and Kini (1994), note that the performance behavior of RLBOs is different from other IPO's since RLBOs have been public before.

The amount of time a company spends in LBO should not be related to performance post IPO as it depends on number of factors including the IPO market conditions and liquidity in the credit markets, that have no specific relation with performance of an individual company. Any

[7] I also checked for any evidence that PE groups renegotiate the time period in which they have to return the capital to the limited partners. I found no news story or academic papers that indicate this. Rather, it seems that many of the PE groups try to return the capital earlier than their contractual time frame.

market-wide or industry-wide effects are already adjusted for in the analyses. On one hand greater amount of time a company spends in LBO may imply that PE group had more time to make beneficial changes and hence lead to better post IPO performance. On other hand, it is also possible that companies that spend more time in LBO do so because they are more problematic and hence need more fixing and hence are expected to have poorer performance post IPO. Recent example of Cerberus's takeover of Chrysler may fall in latter category. Hence ex ante we cannot say for sure that greater time spent in LBO has positive or negative effect on post IPO performance.

I also check for any of the above relations by regressing performance on time spent in LBO, but do not find any significance. Moreover, data reflects that PE groups always decrease their ownership stake at time of the IPO– this further indicates that they are under pressure to sell their stakes.

To test the validity of the instrument econometrically, I regress performance on time since LBO (sum of time spent in LBO plus time since IPO) to catch any mechanical effect of time on performance. As shown in Table III-8 the performance as measured by Tobin's Q is not mechanically related to time since LBO. This gives confidence that at least in this sample the condition of instrument not being related to performance but being related to endogenous covariate is satisfied. However, in case of ROA the instrument is not as strong.

Table III-9 shows the second stage results of panel data regression of performance and PE sponsor ownership. Here PE sponsor ownership is endogenous covariate and the number of years since IPO is the identifying instrument. The Shea partial R-square of first stage regression is 32% and the excluded instrument F-stat is significant at 1% level

indicating that the instrument has good explanatory power and is significant. The high value of R-square ameliorates many of the problems associated with weak instruments.

With ROA as dependent variable the coefficient of PE sponsor ownership is significant while it is significant at 15% level for Tobin's Q, indicating causality between PE sponsor ownership and future performance.

3.5 Conclusion

Using a sample of 133 RLBOs, I find that RLBOs continue to outperform their industry averages up to three years after the IPO year. Dynamics of performance before and after the IPO indicate that PE sponsors conduct the IPO in the year when firms tend to be valued the most. This finding differs from previous literature, where the year of the IPO was also the year when operating performance had peaked. I also find there is deterioration in performance after the IPO. Using cross-sectional, panel and instrumental variable regressions, this paper provides evidence that greater ownership stake by PE sponsors has beneficial impact on future performance. Cross-sectional regression, at time of the RLBO, between performance, and changes in PE sponsor ownership, other insiders' ownership and leverage shows that there is a positive relation between PE sponsor ownership and future performance. I find no relation between ownership of other insiders and leverage with future performance. Since leverage and ownership of other insiders can be endogenous, I do not interpret these results to imply that there is no causal relation between other insiders' ownership and leverage, and performance. However, in case of ownership of PE sponsors, I use the instrument variables approach to ameliorate endogeneity issues and find causality between PE sponsor ownership and future performance. The identifying instrument is the number of years since IPO and is supposed to proxy for impatience of PE sponsors to exit the firms. By comparing the changes in ownership of other insiders with that of PE sponsors, I also find evidence against the

information asymmetry hypothesis. Overall, it appears that greater is the ownership of PE sponsors greater is the beneficial effect on performance of the RLBO firms.

Table III-1. Distribution of LBOs

Distribution of leveraged buyout and their subsequent IPO completion dates by calendar year for 133 firms engaging in reverse LBOs over the period January 1997 to December 2002

Year	Number of LBOs	%	Number of RLBOs	%
1984	1	1%		
1986	1	1%		
1987	2	2%		
1988	1	1%		
1989	2	2%		
1990	2	2%		
1991	2	2%		
1992	2	2%		
1993	2	2%		
1994	3	2%		
1995	12	9%		
1996	31	23%		
1997	24	18%	23	17%
1998	19	14%	21	16%
1999	17	13%	26	20%
2000	9	7%	26	20%
2001	3	2%	20	15%
2002	0	0%	17	13%
Total	133	100%	133	100%

Table III-2. Industry Distribution

Industry distribution of firms that completed RLBO over the period January 1997 to December 2002.

SIC Codes	Industry Description	Number Of RLBO Firms	%
0-1000	Food products	0	0%
1000-1999	Mining and construction	8	6%
2000-2999	Consumer Products	15	11%
3000-3999	Manufacturing	31	23%
4000-4999	Utilities and transportation	19	14%
5000-5999	Wholesale, retail and some services	21	16%
6000-6999	Financial Services	6	5%
7000-7999	Personal & business services	18	14%
8000-8999	Miscellaneous	15	11%
Total		133	100%

Table III-3. Summary Statistics

Summary statistics in the year of IPO of firms that completed RLBO over the period January 1997 to December 2002. Firms in utility and financial categories were excluded from the original sample of 133 firms.

Variable	Obs.	Median	Mean	Std. Dev.	Min.	Max.
Return On Assets	116	0.021	-0.379	2.459	-25.188	0.351
Return On Sales	116	0.021	-0.005	0.167	-0.848	0.322
Capex/PPE	116	0.325	0.358	0.215	0.063	0.978
R&D/Assets	116	0.000	0.030	0.071	0.000	0.391
Adv/Assets	116	0.000	0.016	0.035	0.000	0.159
CashHoldings/PPE	116	0.211	2.474	7.811	0.000	72.979
Leverage	116	0.502	0.533	0.574	-0.083	4.203
Tobin's Q	115	1.826	2.492	2.239	0.544	13.33
Sales	116	236.72	585.28	968.30	0.72	6,772.8
Assets	116	313.52	717.99	1,821.65	8.06	18,966.5

Table III-4. Performance of RLBO companies

Industry Adjusted Performance of 116 RLBO companies. Benchmark is the mean industry performance for the corresponding year. IPO+0 is the year of RLBO. Significance levels are calculated using Wilcoxon signed-rank test. One, two and three stars imply significance at 10%, 5% and 1% respectively.

Year	Tobin's Q	ROA	ROS	Leverage	Number Of Obs.
IPO-2		0.031***	0.032***	0.245*	82
IPO-1		0.021***	0.035***	0.352*	110
IPO+0	0.186***	0.048***	0.051***	0.042	116
IPO+1	0.169***	0.072***	0.059***	-0.021	112
IPO+2	0.109***	0.066***	0.051***	-0.066	106
IPO+3	0.015	0.039***	0.037***	-0.031	100
IPO+4	-0.146	0.011	0.021	-0.052	90

Table III-5. Summary Statistics

Percent ownership of private equity sponsors and other insiders for 73 RLBOs between 1997 and 2002 that had ownership data in the proxy statements. Management and other insiders include all officers and directors who are not representatives of the PE firms. Ownership is defined as beneficial ownership and includes the shares that can be issued by exercise of options or warrants within 60 days of the date of proxy statement.

Panel A: Private Equity Investors Percent Ownership

Year	Mean	Std. Dev.	Min.	Max.
IPO-1	54.1	24.3	8.2	96.3
IPO+0	38.4	18.1	1.7	84.1
IPO+1	35.6	18.6	0	81
IPO+2	30.6	20.6	0	81
IPO+3	28.1	20.5	0	76.2
IPO+4	23.6	20.8	0	69.2
IPO+5	19.2	19.1	0	57.2
IPO+6	20.3	17.5	0	55
IPO+7	13.2	17.4	0	54.3

Panel B: Management and Other Insiders Percent Ownership (excluding PE sponsors)

Year	Mean	Std. Dev.	Min	Max
IPO-1	21.9	18.7	0.7	61.7
IPO+0	17	14.4	0.7	57.5
IPO+1	15.4	13.1	0.6	56.7
IPO+2	13	10.9	0.9	49.6
IPO+3	12.4	10.7	0.8	41.7
IPO+4	11.9	11.1	0.3	44.4
IPO+5	11.8	11.7	0.3	41
IPO+6	8.8	9.7	0.5	38.6
IPO+7	7.6	7.5	1.3	25.8

Table III-6. Cross-Sectional Regression of Performance on Ownership and Leverage

Following table shows the results of the cross-section regression with following specification:

$$AvgQ_i = \alpha + \beta_1 ChngPEOwn_{i,\,(IPO-1\ to\ IPO+1)} + \beta_2 ChngMgmtOwn_{i,\,(IPO-1\ to\ IPO+1)} + \beta_3 ChngLev_{i,(IPO-1\ to\ IPO+1)} + \beta_4 QatIPO_i + \beta_5 Reputation_i + \varepsilon_i \text{ , where } i=1,2,.....,N.$$

AvgQ is averaged over two, three and four years after the IPO and is proxy for long term performance. Tobin's Q and leverage are industry adjusted. ChangePEOwnership is change in absolute percentage ownership from immediately before IPO to immediately after the IPO, for example, if PE sponsors own 50% of firm pre-IPO and own 40% post-IPOthe change will be recorded as negative 10%. ChngMgmtOwn is change in absolute percentage ownership from immediately before IPO to immediately after the IPO by all insiders(officers and directors) who are not affiliated with private equity firm. t stats are reported in parenthesis and are based on robust standard errors. *, **, *** indicates significance at the 10%, 5% and 1% level, respectively.

	Average Q(IPO+1 to IPO+2)	Average Q(IPO+1 to IPO+3)	Average Q(IPO+1 to IPO+4)	Average ROA(IPO+1 to IPO+2)
ChangePEOwnership	0.0166**	0.0136***	0.0112**	0.0007
	(2.42)	(2.52)	(2.24)	(-0.92)
ChangeMgmtOwnership	0.001	0.0012	0.0018	-0.0001
	(0.24)	(0.39)	(0.73)	(-0.26)
ChangeLeverage	-0.0194	0.0019	0.0238	-0.008
	(-0.40)	(0.04)	(0.53)	(-0.92)
PerfAtIPO	0.459***	0.376***	0.329***	0.165
	(3.15)	(2.92)	(2.85)	1.16
ReputationDummy	-0.012	0.01	0.008	-.023
	(-0.06)	(0.06)	(0.05)	(-0.85)
R-Square	0.49	0.46	0.45	0.09
Number of Observations	73	73	73	73

Table III-7. Panel Data Regression of Performance on Ownership and Leverage(Top quartile RLBOs)

Following table shows the results of the cross-section regression with following specification:
$AvgQ_i = \alpha + \beta_1 ChngPEOwn_{i,\ (IPO\text{-}1\ to\ IPO+1)} + \beta_2 ChngMgmtOwn_{i,\ (IPO\text{-}1\ to\ IPO+1)} + \beta_3\ ChngLev_{i,(IPO\text{-}1\ to\ IPO+1)} + \beta_4 QatIPO_i + \beta_5 Reputation_i + \varepsilon_i\ ,\ where\ i=1,2,.....,N.$
AvgQ is averaged over two, three and four years after the IPO and is proxy for long term performance. Tobin's Q and leverage are industry adjusted. ChangePEOwnership is change in absolute percentage ownership from immediately before IPO to immediately after the IPO, for example, if PE sponsors own 50% of firm pre-IPO and own 40% post-IPOthe change will be recorded as negative 10%. ChngMgmtOwn is change in absolute percentage ownership from immediately before IPO to immediately after the IPO by all insiders(officers and directors) who are not affiliated with private equity firm. t stats are reported in parenthesis and are based on robust standard errors. *, **, *** indicates significance at the 10%, 5% and 1% level, respectively.

	Average Q(IPO+1 to IPO+2)	Average Q(IPO+1 to IPO+3)	Average Q(IPO+1 to IPO+4)
ChangePEOwnership	0.0164**	0.0126***	0.008
	(2.38)	(2.34)	(1.50)
ChangeMgmtOwnership	-0.007	-0.0043	-0.0003
	(0.24)	(-1.29)	(-0.10)
ChangeLeverage	-0.113	-0.075*	-0.064**
	(-1.71)	(-1.82)	(-2.53)
PerfAtIPO	0.922***	0.676***	0.52***
	(9.57)	(10.59)	(13.85)
ReputationDummy	-0.41	0.43	0.38
	(1.32)	(1.58)	(1.51)
R-Square	0.92	0.90	0.87
Number of Observations	18	18	18

Table III-8. Panel Data Regression of Performance on Ownership and Leverage

Following table shows the results of the panel data regression with following specification:

$Perf_{i,t+1,\,t+2} = \alpha_i + \beta_1 PEOwn_{i,t} + \beta_2 MgmtOwn_{i,t} + \beta_3 Lev_{i,t} + \beta_4 LogSales_{i,t+2} + \varepsilon_{i,t+1,t+2}$ *, where t= IPO+0, IPO+1,.....,IPO+6 and i=1,2,.....,N.*

PEown is percentage of shares held by PE sponsors at time t. AvgPerf is the average of performance in the years t+1 to t+2. Performance and leverage(Lev) are industry adjusted. t stats are reported in parenthesis and are based on robust standard errors. *, **, *** indicates significance at the 10%, 5% and 1% level, respectively.

	Tobin's Q	ROA	Tobin's Q	ROA
PEOwnership	0.0123**	0.0016***	0.0130**	0.0007
	(2.44)	(3.63)	(2.42)	(1.25)
MgmtOwnership	0.0077	0.0018	0.0082	0.0011
	(1.43)	(1.64)	(1.48)	(1.10)
Leverage	-.046	-0.0095	-0.044	-0.012
	(-0.74)	(-1.47)	(-0.70)	(-1.85)
LogSales	-.0127	0.077***	-0.003	0.0646**
	(-.05)	(2.44)	(-0.01)	(2.10)
TimeSinceLBO			0.0096	0.0122***
			(0.25)	(-2.61)
R-Square	0.068	0.12	0.068	0.15
Firm Fixed Effects	Yes	Yes	Yes	Yes
Number of Observations	241	241	241	241

Table III-9. 2SLS IV Regression of Performance on Ownership and Leverage

Following table shows the results of the second stage panel data regression with following specification:

Stage I: $PEOwn_{i,t} = \alpha_i + \gamma_1 TimeSinceLBO_{i,t} + \gamma_2 MgmtOwn_{i,t} + \gamma_3 Lev_{i,t} + \gamma_4 LogSales_{i,t+2} + \varepsilon_{i,t}$

Stage II: $AvgPerf_{i,t+1,\,t+2} = \alpha + \beta_1 PEOwn_{i,t} + \beta_2 MgmtOwn_{i,t} + \beta_3 Lev_{i,t} + \beta_4 LogSales_{i,t+2} + \varepsilon_{i,t+1,t+2}$,

where t= IPO+0, IPO+1,.....,IPO+6 and i=1,2,.....,N.

TimesinceLBO is number of years since LBO. PEown is percentage of shares held by PE sponsors at time t. AvgPerf is the average of performance in years t+1 to t+2. Performance and leverage are industry adjusted. t stats are reported in parenthesis. *, **, *** indicates significance at the 10%,5% and 1% level, respectively.

First Stage

	PEOwnership	
MgmtOwnership	-0.1308	(-1.34)
Leverage	0.2453	(0.21)
LogSales	-0.2665	(-0.07)
TimeSinceLBO	-4.038	(-8.79)
R-square First Stage	0.33	
Shea Partial R-square	0.318	
Instrumental F-Stat	30.52	
Firm Fixed Effects	Yes	
Number of Observations	241	

Second Stage

	Tobin's Q		ROA	
PEOwnership	0.0107	(1.49)	0.0037***	(3.51)
MgmtOwnership	0.0079	(1.34)	0.0015*	(1.75)
Leverage	-0.043	(-0.58)	-0.0126	(-1.15)
LogSales	-0.0035	(-0.01)	0.0655*	(1.86)
R-square	0.07		0.05	
Firm Fixed Effects	Yes		Yes	
Number of Observations	241		241	

Appendix III-1: Variable Definitions

Accounting and ownership variables used in the paper are defined as follows:

Return on assets is the ratio of EBITDA (Compustat item 18) over total assets (Compustat item 6).

Return on sales is the ratio of EBITDA (Compustat item 18) over total sales (Compustat item 12).

Tobin's Q is defined as the market value of assets divided by the book value of assets (Compustat item 6). The market value of assets equals the book value of assets plus the market value of common equity (common shares outstanding (Compustat item 25) times fiscal year end price (Compustat item 199)) less the sum of the book value of common equity (Compustat item 60) and balance sheet deferred taxes (Compustat item 74).

Leverage is long-term debt (Compustat item 9) plus current liabilities (Compustat item 34) divided by long-term debt plus current liabilities plus book value of common equity (Compustat item 60).

PEOwnership is the absolute percentage, beneficial ownership of all officers or directors who are employees or partners of the private equity firm. In calculation of percent ownership, I include shares of common stock that can be issued upon the exercise of outstanding options and warrants that are currently exercisable or exercisable within 60 days of the date of proxy statement.

Mgmt ownership or Other insiders ownership is the absolute percentage, beneficial ownership of all officers or directors excluding the PE sponsors. In calculation of percent ownership, I include shares of common stock that can be issued upon the exercise of outstanding options and warrants that are currently exercisable or exercisable within 60 days of the date of proxy statement.

Reputation Dummy –This variable is 0 or 1 based on low or high reputation of PE firm. I searched for each PE firm to find its age and funds under management. For two of the companies I was not able to get any information. These two firms were categorized as low reputation firms. As proxy for reputation I use the size of the fund and its age. Lerner, Schoar 2004 (JFE) find that bigger and older PE funds tend to be more successful. A fund that has generated success in the past is expected to continue and raise more capital.

Hence I assume that bigger (older) fund had more series of successes and hence is more reputable. This proxy for reputation is used in the cross section regressions at time of the IPO. For panel data I do not need to adjust for reputation as firm fixed effects cancel out any effect of reputation.

Appendix III-2: Examples of unobserved firm heterogeneity effecting performance and ownership

Himmelberg, Hubbard, & Palia (1999) detail three examples of unobserved heterogeneity that can induce positive correlation between ownership and performance. Using a panel data specification can ameliorate the endogeneity concerns if we assume that unobserved heterogeneity is a firm fixed effect

Example I - two firms are identical except that one has higher proportion of intangible assets than other. Optimal contract for the firm with higher proportion of intangible assets will involve higher lever of ownership as intangible assets are hard to monitor. But at same time firm will have higher Q value by definition. Hence, the proportion of intangibles induces a positive correlation between ownership and performance.

Example II - two firms are identical except that one has access to superior monitoring technology. Optimal contract for the firm with superior monitoring technology will require lower lever of ownership to align incentives and at same time will have higher valuation because fewer resources are diverted to managerial perquisites. Hence, the presence of superior monitoring technology induces a correlation between ownership and performance.

Example III - two firms are identical except that one has greater market power that insulates its managerial decision-making from the discipline of product markets. Optimal contract for such firm will require higher level of ownership to compensate for lack of monitoring by product markets. At same time greater market power will translate into higher Q. Hence, the difference in market power between two firms induces a positive correlation between ownership and performance.

References

Risk Reduction as a Managerial Motive for Conglomerate Mergers, Yakov Amihud and Baruch Lev , *The Bell Journal of Economics*, Vol. 12, No. 2 (Autumn, 1981), pp. 605-617

Bharat A. Jain and Omesh Kini,1994, The Post-Issue Operating Performance of IPO Firms, *The Journal of Finance*, Vol. 49, No. 5 (Dec., 1994), pp. 1699-1726

Cao, Jerry X., "What Role Does Private Equity Play When Leveraged Buyouts Go Public?" (February 20, 2008, SSRN working paper)

Cao, Jerry X., and Joshua Lerner, 2007, the Performance of Reverse Leveraged Buyouts, *NBER Working Paper*

Chou, De-Wai, Michael Gombola, and Feng-Ying Liu, 2006, Earnings Management and Stock Performance of Reverse Leveraged Buyouts*, Journal of Financial and Quantitative Analysis,* forthcoming

Degeorge, Francois; Zeckhauser, Richard, 1993, "The Reverse LBO Decision and Firm Performance: Theory and Evidence," *Journal of Finance*, 48, 1323-48

Eugene F. Fama , Agency Problems and the Theory of the Firm, *The Journal of Political Economy*, Vol. 88, No. 2 (Apr., 1980), pp. 288-307

Eugene Fama and Michael Jensen, Separation of ownership and control, *Journal of Law and Economics* 26 (1983), pp. 301–325.

Eugene Fama and Michael Jensen, Organizational forms and investment decisions, *Journal of Financial Economics* 14 (1985), pp. 101–120.

Himmelberg, Charles P. & Hubbard, R. Glenn & Palia, Darius, 1999. "Understanding the determinants of managerial ownership and the link between ownership and performance," *Journal of Financial Economics*, Elsevier, vol. 53(3), pages 353-384.

Holthausen, Robert W., and David F. Larcker, 1996, "The Financial Performance of Reverse Leveraged Buyouts," *Journal of Financial Economics*, 42, 293-332.

Kaplan, Steven N., 1989, "The Effects of Management Buyouts on Operating Performance and Value," *Journal of Financial Economics*, 24, 217-254.

Kaplan, Steven N., 1997, "The Evolution of U.S. Corporate Governance: We are All Henry Kravis Now," *Journal of Private Equity*, 1, Fall, 7-14.

Kaplan, Steven N., and Jeremy C. Stein, 1993, "The Evolution of Buyout Pricing and Financial Structure in the 1980s," *Quarterly Journal of Economics*, 108, 313-357.

Leland, H. E., and D. Pyle, 1977, Information asymmetries, financial structure, and financial intermediation, *Journal of Finance* 32, 371-387.

Lerner, Joshua, and Antoinette Schoar, 2004, The illiquidity puzzle: Theory and evidence from private equity*, Journal of Financial Economics* 72, 3– 40

Lerner, Joshua,, Antoinette Schoar, Wan Wongsunwai (2007) Smart Institutions, Foolish Choices: The Limited Partner Performance Puzzle *The Journal of Finance* 62 (2) , 731–764

Mian, Shehzad, and James Rosenfeld, 1993, "Takeover Activity and the Long-Run Performance of Reverse Leveraged Buyouts," *Financial Management*, 22, Winter, 46-57

Michael Jensen, The eclipse of the public corporation, *Harvard Business Review* 67 (1989), pp. 61–74.

Michael Jensen and William Meckling, Theory of the firm, managerial behavior, agency costs and ownership structure, *Journal of Financial Economics* 3 (1976), pp. 305–360

Muscarella, Chris J., and Michael R. Vetsuypens, 1990, "Efficiency and Organizational Structure: A Study of Reverse LBOs," *Journal of Finance*, 45, 1389-1413.

Wayne H. Mikkelson, M. Megan Partch, Kshitij Shah, Ownership and operating performance of companies that go public, *Journal of Financial Economics* Volume 44, Issue 3, , June 1997, Pages 281-307

Zingales, Luigi and Novaes, Walter, "Capital Structure Choice When Managers Are In Control: Entrenchment versus Efficiency" (December 1995). NBER Working Paper No. W5384