The Predictability of Analyst Coverage on Stock Returns

- Empirical Evidence from China's Stock Market

Zhuoyan Lao[1] and Huaqian Zhu[2]

Master Thesis in Finance

Stockholm School of Economics

Department of Finance

Tutor: Michael Halling

Spring 2018

[1] 41023@student.hhs.se
[2] 41005@student.hhs.se

Abstract

This paper studies the association between the analyst coverage (both total analyst coverage and abnormal analyst coverage) and future stock returns in Shanghai A-share stock Market over a period of ten years from 2008 to 2017. Our study draws inspiration from the work of Charles M.C. Lee and Eric C. So (2016). We first get the abnormal analyst coverage by decomposing observed analyst coverage into expected and abnormal parts applying residual analyst coverage model. Then by applying portfolio sorts and Fama-Macbeth panel regression, we find that abnormal analyst coverage, which is unobserved in the market, is positively associated with future stock returns indicating that stocks that receive abnormally higher (lower) coverage from analysts are followed by higher (lower) returns. Based on this finding, we further prove that a monthly-rebalanced strategy that longs stocks in the highest quintile of abnormal analyst coverage and at the same time shorts stocks included in the lowest quintile could provide an annualized return of approximately 10% on average, robust to standard asset pricing factors (RMRF, SMB, HML, MOM). These findings have great significance on the study of the predictability of analyst coverage, especially in Chinese market, as well as on the guidance of investing in China's stock market.

Keywords: Analyst Coverage, Future Stock Return, China's Stock Market

Acknowledgements: We would like to express our sincere gratitude to our tutor Michael Halling for valuable input and guidance throughout the entire thesis writing process.

Table of Contents

1. Introduction

In this part, we first present an overview of the development of both China's stock market and analysts in China. Then we discuss the motivation and contribution of our study in order to show a full picture of the whole study. An outline of this paper is presented at the end of this part.

1.1 Background

Over the past 27 years since its born in 1990, China stock market has been through many fluctuations and has witnessed the rapid development of Chinese economy. We can see the impacts that big events happened in the process of Chinese economy development including the reform of ownership structure of state-owned shares, the 4 trillion economic stimulus plan, One Belt and One Road Initiative, etc. exerted on China stock market. Sharing a common fate with Chinese economy and China stock market, analysts in China has grew a lot, from the perspective of both volume and experience though there are still many issues Chinese security analysts are facing such as balancing speed and quality.

1.1.1 The development of China's stock market

China's stock market was born at the early stages of economic reforms in which the economic system was slowly transitioning from a planned economy to a market economy. Chinese stock market is one of a kind. The fact that China was a Communist Republic did not prevent it from establishing functioning capital markets. Since the re-establishment of the stock exchange in the 1990s, the market has been through some huge changes. Since the SME-Board, NEEQ and ChiNext has established, it has become much easier for smaller companies to float shares, which stimulates the growth of these young ventures. The trend towards internationalization and the steady rise in disposable income in the past 20 years have enabled China to become the second largest equity market in the world. However, it's still too early to call the Chinese stock market a truly global open marketplace. Investors, both foreign and domestic investors, individual and institutional investors, still face a lot of concerns. First and foremost, the State Council of the PRC (People's Republic of China) and the CSRC (China Securities Regulatory Commission) still have influential power on the capital market. The fact that they can intervene the market at a time they believe necessary still scares off a lot of investors although this behavior sometimes plays as a positive role in keeping the market stable for a period of time. Secondly, the Chinese government and all of its affiliated organizations still hold a fairly large part of shares. This creates governance problems and affects investors as well though the proportion of government and its affiliated organizations holding shares is decreasing. Furthermore, information disclosure often remains vague and individual shareholder protection deserves more attention from regulators. China A-share stock index has increased by three times from around 1000 points to 3000 points over the past 20 years. However, during these 20 years, A-share market has been through some big fluctuations as well. From May,1999 to June 1999, A-share index increased 67.7% from 1047 points to 1756 points. Then the market dropped to 1341 points in the six months after the launch of *Securities Law* in July 1999. Later on, the market started to recovery and it increased to 2245 points by the end of June 2001. However, with the announcement of state-owned shareholding reduction, A-share market started to fall down and it dropped to 998.2 points after years in June, 2005. Since then, SEC came up with a series of policy with the purpose to stimulate stock trading including *the Reform Scheme of Ownership Structure of State-owned Shares*, which gave a portion of state-owned shares to outstanding shareholders. The market finally recovered and the famous bull market in 2006 and 2007 happened. By the end of October in 2007, A-share index reached 6125 points. Realized that there was a huge bubble in China stock market, the government started to suppress the market. The policies included hikes in the lending rates and banks reserve requirement ratio. Together with the influence of the financial crisis in the U.S., A-share stock market has been through the hardest time in 2008. By the end of October in 2008, A-share index went back to 1664 points. After the huge decrease, on November 11th, 2008, Chinese government announced the famous Chinese economic stimulus plan, which is a RMB¥ 4 trillion (US$586 billion) stimulus package as an attempt to minimize the impact of the global financial crisis on the world's second largest economy. Later in 2009, stock market reacted to this plan and the A-share index increased again. However, 6000 points never happened again and on August 4th in 2009, market index reached 3478 points. In the following 4 years since 2010, the stock market fell down slowly in general despite of the good performance of stocks listed on the growth enterprise market which opened in October 2009. The recent bull market started from July 2014 and ended up in June 2015. When the market index dropped to 2000 points in 2014, a large amount of investors entered the market and bought at bottom. Together with the announcement of "One Belt and One Road Initiative" and the implementation of state-owned enterprise reform, the market increased again and reached 5178 points in 2015. Since then, A-share market started to drop at a lower rate with small fluctuations and ended up with 3247 points by the end of April, 2018[3].

1.1.2 Analysts in China

China's securities analyst industry has gone through more than 30 years of development along with the growth of China's securities market. At present, China's securities analysts are facing the challenge of fierce competition among both domestic analysts and foreign analysts since China’s stock market, as a promising emerging market, is getting more attention from all over the world.

According to the implementation process of the supervision from Chinese government, the development of analysts in China can be divided into three stages.

At the first stage (1984-1991), which is called initial stage, the Chinese stock market has not yet officially established. All parties in the market including listed companies, securities companies and investors were crossing the river for the first time. A few people began to study the knowledge about capital market, but a stable securities and consulting companies has not yet formed.

At the second stage (1991-1998), securities analyst industry grew rapidly in volume. When China’s stock market was found in the early 1990s, securities companies started to establish securities consulting department in order to meet the needs of securities consulting from investors. But at that time, most of the analysts did not have sophisticated knowledge and experience. It would be more accurate to call the service they provided an introduction to stock market. The absence of supervision from government speeded up the development of both China’s stock market and analysts in China during that period.

At the third stage (1998-present), with securities regulations and laws for the supervision and administration entering into the market, China’s securities market together with securities firms started to be on the right track. At the same time, analysts started to focus more on professional skills, such as selecting stock, conducting research, etc.

After 18 years since 1998, the number of registered analysts in 95 securities companies reached 2283 at the end of 2016[4]. Analysts now are under too much pressure of competition due to the rapid growth of China’s securities market. A direct and negative effect of the fierce

[3]All the A-share composite index mentioned in 1.1.1 could be found in Wind.
[4] Data source: Securities Association of China (SAC). See on http://www.sac.net.cn

career competition in the Chinese analyst industry is that it compels analysts to issue reports frequently and quickly, sometimes leading to poor quality analyses.

1.2 Motivation

As essential information intermediaries in global capital markets, analysts are supposed to facilitate information dissemination in capital markets, and to add value for their clients by helping investors identify and select good investment opportunities to earn satisfying returns in the foreseeable future. Many previous studies (Finegan et al. 1996; McNichols and O'Brien. 1997; Gasparino and Smith. 2002; Somnath. 2006; Lee and So. 2016) have documented that analysts tend to avoid following companies about which they hold unfavorable opinions and to hold back negative news, while to cover firms with favorable expectations. Such selective coverage by analysts might be the result of many economic disincentives and conflicts of interest for analysts[5] (see McNichols and O'Brien. 1997; Gasparino and Smith. 2002; Somnath. 2006 for detailed disincentives). Researchers such as Finegan et al (1996), Lee and So (2016) also explain this phenomenon using the framework of allocation behavior of individuals given resource constraints and expected payoff. The most common hypothesis researchers have made based on analysts' selective coverage is that observed analyst coverage contains information about future firm performance measured by future stock returns. In other words, the observed number of analysts covering a firm should be a robust return predictor, if analysts do have the ability to pick high quality firms with promising prospects and tend to selectively cover these glamour firms rather than unfavorable firms. Analysts' buy/sell/hold recommendations about those firms they choose to cover are also considered to be linked with future stock return.

Many studies focus on analyst recommendations and find a significant link between analyst recommendations and future stock returns. Jerring (1983) shows that investing in analyst recommended stocks can earn significant positive abnormal profit after deducting transaction costs, but the stocks recommended by analysts in his study are densely distributed in the natural resources related industry. Trueman et al (2002) also finds a significant positive abnormal profit by longing high-rating recommended stock at the same time shorting low-rating stocks, but after deducting related transaction costs the profit become insignificant. Womack (1996) presents a significant relation between analysts' buy-sell recommendations and short-term stock price/return. Lee et al. (2001) further discover that the predictive power

[2]In the section of literature review, you can find more details about those economic disincentives and conflicts of interests that are considered to lead to selective analyst coverage.

of recommendation changes (revisions) is more robust than the level of recommendations. Oya (2016) examines the post-revision drift following recommendation changes and reports that upgrades (downgrades) are followed by positive (negative) return.

By contrast, empirical tests on the predictive ability of analyst coverage on future return are far less in numbers and reached to controversial conclusions. Somnath et al (2006) mainly put their eyes on the new issue market and find that in the three years subsequent to initial coverage, the return of IPOs with high residual coverage[6] lead the return of those with lower residual coverage. But such positive relation is conditioning on specific firm event, IPO. Ambrus et al (2006) document a negative relation between analyst coverage and future returns, showing that future return of a firm will be lower when the total number of analysts covering the firm increases, and higher when analyst coverage decreases, whereas Jung (2015) reports that increases in analyst interest are positively related to stock returns over the next three months. Wang and Yao (2008) apply the residual coverage model used by Somnath to examine the relationship of stock return and analyst coverage in the Shanghai Exchange A-share market of China. But they manually collect only one-year analyst coverage data as their sample. So, their results are questionable. Recently some progress has been made concerning this specific topic. Lee and So (2016) develop a broadly applicable approach to extract future return information from analyst coverage data, not conditioning on any relatively rare firm events, and show that in developed markets, to be specific, the US stock market, firms with high abnormal/residual coverage outperform firms with low abnormal/residual coverage in subsequent months. They also prove that part of the return predictability stems from analyst increasing abnormal coverage in underpriced stocks. But this approach has never been applied in analysis of analyst coverage in emerging markets.

Actually, so far most empirical studies about selective analyst coverage and the predictive power of analyst coverage or recommendations on future returns focus on developed markets (Jerring, 1983; Finegan et al. 1996; Womack, 1996; McNichols and O'Brien. 1997; Lee et al. 2001; Gasparino and Smith. 2002; Jung, 2015; Somnath. 2006; Lee and So. 2016). Few researches studies emerging markets (Hameed et al. 2004, studied global emerging markets; Wang and Yao, 2008, Chinese market; Francisco Marcet, 2017, Latin American market).

---

[6] Residual analyst coverage (Hong et al. 2000; Somnath et al. 2006; Wang and Yao. 2008) or abnormal analyst coverage (Lee and So. 2016) referred to the portion of analyst coverage that cannot be explained by commonly known determinants of analyst coverage such as firm size, share turnover, past performance, etc. That is, the residuals from the regression of total analyst coverage on commonly known factors.

That situation motivates us to revisit the question of whether analyst coverage can be a return predictor and to test the hypothesis that analyst coverage contains some information about future stock returns (that is, it has a significant relationship with future stock returns) in Chinese Shanghai Exchange A-share market, a representative fast-growing and emerging market. And we further hypothesize that the relation between analyst coverage and future returns should be positive. Our hypothesis are consistent with and based on 1) relevant literature on analyst selective coverage for firms about which they have favorable expectations either because of economic incentives or resource limits (O'Brien, 1997; Lee and So, 2016.); 2) literature about analyst coverage imposing significant positive influence on firms' future market value and cost of capital (Kee and Jo, 1996; Easley and O'Hara, 2004); 3) empirical evidence showing that abnormal/residual analyst coverage is positively correlated with future stock returns in new issue market (Somnath, 2006) and in developed market (Lee and So, 2016).

1.3 Contributions

Based on the hypothesis that 1) analysts cautiously select the firms they cover relying on their expectations about firms' future performance without consideration of maintaining good relationship with listed firms; 2) analyst coverage contains information about future firm performance measured by future stock returns; 3) abnormal analyst coverage, is positively correlated with future stock returns, we 1) apply the residual analyst coverage model developed and improved by Hong et al. (2001), Sonmath et al. (2006), Wang and Yao (2008), Lee and So (2016), to decompose observed total analyst coverage into expected and abnormal parts and modify the model by adding ROA as another factor. 2) implement Portfolio sorts and Fama-Macbeth panel regression to explore the relationship between future stock returns and both total analyst coverage and abnormal analyst coverage respectively; 3) apply Fama four factor asset pricing model to test the robustness of the returns from strategy exploiting the link between analyst coverage and future share returns.

The results show that 1) abnormal analyst coverage is positively associated with future stock returns after excluding outliers in Dec 2008, indicating that stocks that receive abnormally higher (lower) coverage from analyst are followed by higher (lower) returns, and the monthly-rebalanced strategy that longs stocks in the highest quintile of abnormal analyst coverage at the same time shorts stocks included in the lowest quintile earns an annualized return of approximately 10% on average, robust to standard asset pricing factors (RMRF, SMB, HML, MOM). Thus, abnormal analyst coverage can be a powerful return predictor.

These conclusions contribute to the studies focusing on analysts' behavior, analyst coverage and the association between analyst coverage and stock future returns in Chinese market. The main contribution of this study is that we supplement the empirical evidence of the positive link between abnormal analyst coverage and future stock returns testing the data over the past ten years from 2008 to 2017 in Chinese market. In addition, based on this positive relation, we further prove that the investing strategy of longing stocks in the highest quintile of abnormal analyst coverage and at the same time shorting stocks included in the lowest quintile could provide promising return for investors. Moreover, the decomposition of total analyst coverage into four factors including Size, Share Turnover, Momentum and ROA, especially the first time of introducing ROA into the model, provides a better understanding of the determinants of total analyst coverage in Shanghai A-share stock market.

1.4 Outlines

This paper first introduces the background of China's stock market and the development of and security analysts in China as well as our motivation to study on the association between analyst coverage and future stock returns in Chinese market and our contribution to this topic. In the second part, we review former studies on analyst coverage, stock price informativeness and the association between analyst coverage and stock future returns both in foreign markets and in Chinese markets. In the third part, we first elaborate the methodologies we apply in this study, which are Abnormal/Residual Analyst Coverage Model, Portfolio Sorts, Fama-Macbeth Panel Regression and Fama Four Factors Model. Then we present our data sources, data cleaning process and elaborate the data description. In the fourth part, we show the detail of our empirical analysis by elaborating every result involving in the model/approach/method we apply. And in the last part, we provide our conclusions that abnormal analyst coverage is positively associated with future stock returns in Shanghai A-share stock market and discuss the contributions as well as the issues we have not covered that further study is still needed to conduct on.

2 Literature Review

In this part, we review former studies on analyst coverage, stock price informativeness and the links between analyst coverage and future stock returns both in foreign markets and in Chinese markets.

2.1 Analyst Coverage

The earliest study on the determinants of the analyst coverage is from Ravi Bhushan (1989). He examines the major determinants of the number of analysts following a firm. A simple model of analyst following is proposed and several firm characteristics are suggested that are likely to influence the extent of a firm's analyst following by either affecting the aggregate demand for or supply of analyst services or both for the firm. Specifically, the level of analyst following for a firm is positively associated with level of institutional investment, variability of returns, the correlation between the firm return and the market return and firm size. There is a negative association with the level of insider shareholdings and industrial diversification. Almost all of these characteristics are found to be strongly significant in affecting the extent of analyst following of firms and the empirical results generally accord well with economic intuition.

Maureen McNichols and Patricia O'Brien (1997) further examines the relation between analysts' information about a stock's future prospects and their decisions to issue investment recommendations and earnings forecasts for that stock, and the implications of this relation for the observed distribution of recommendations and earnings forecast errors. Based on the articles that have argued that analysts are reluctant to issue unfavorable investment information, perhaps because they fear jeopardizing potential investment banking business;1: they fear losing access to management as a source of information; 2: and/or they seek to generate trading commissions.3: these forces cause analysts to bias their true predictions toward a more optimistic view, they examine an alternative response to disincentives to disclose negative information, that analysts are more likely to provide forecasts and recommendations for stocks about which their true expectations are favorable.

Mark T. Bradshaw (2002) studies on the justifications for analysts’ stock recommendation. He studies on 103 security analysts' reports, specifically, extract the frequency that analysts use target prices as justifications for their stock recommendations. Moreover, he further examines the association between the extent of observed overpricing or underpricing indicated in target prices and the favorableness of stock recommendations. The results show that analysts disclose target price justifications in over two-thirds of the sample reports, and at the same time target prices show a positive relation with the favorableness of stock recommendations. Price-to-earnings ratios and expected growth have higher frequency to be used to justify the most favorable recommendations (and target prices), while other qualitative statements are more likely to be used to justify the least favorable recommendations. He also finds that analysts usually compute target prices applying price-multiple heuristics such as price-earnings-to-growth, "PEG". Different explanations are proposed, including self-selection biases suggesting that when analysts are less confident about underlying earnings forecasts, they tend not to disclose target prices.

Later in 2004, Mark further examines whether valuation estimates based on analysts' earnings forecasts are consistent with their stock recommendations. Because earnings forecasts are linked to value and recommendations reflect analysts' opinions of value relative to current price, earnings forecasts and stock recommendations should be linked in a predictable manner. He considers four possible valuation models of how earnings forecasts and stock recommendations are linked. These models include two specifications of the residual income model, a price-earnings-to-growth (PEG) model, and analysts' projections of long-term earnings growth. The results provide little evidence that analysts' recommendations are explained by either residual income model specification. However, both the PEG model and analysts' projections of long-term earnings growth explain analysts' stock recommendations. The relation between the valuation models and future returns is also examined. Analysts' projections of long-term earnings growth have the greatest explanatory power for stock recommendations, but investment strategies based on these projections have the least association with future excess returns. Overall, this study suggests that analysts' recommendations are more correlated with heuristic valuation models than with present value models and buy-and-hold investors would earn higher returns relying on present value models that incorporate analysts' earnings forecasts than on analysts' recommendations.

In Chinese stock market, Song Zhu, Xiaoyu Jiang and Xiaoli Ke (2016) investigates the impact of stock index adjustment on analyst coverage. Their results show that the stock index adjustment exerts a significant effect on analyst coverage after studying on 231 pairs of matched firms from year 2009 to year 2012 in Chinese stock market. Adding to the stock index can exert more analyst coverage. In Contrast, deleting from the stock index dose not have significant effect, which shows an implication that stock index adjustment has an significant impact on the information environments of firms that have been added into the stock index. An index adjustment also has impacts on institutional holdings with the consideration of new information such as changes in fundamentals and information environments. Changes in institutional shareholdings are partially attributable to the changes in analyst coverage, and both index funds and other types of funds can change their portfolios in response to the changes in the target firms' informativeness.

Huai-Chun Lo and Febri Rahadi (2017) examines the association between risk-adjusted returns and analyst coverage. Their results show that analysts are more likely to be follow and write reports on stocks with better risk-adjusted returns. Following and studying on these stocks could be an efficient and effective way to provide high quality reports in a shorter time. Further evidence show that earnings forecasts provided by analysts are more accurate for stocks with better risk-adjusted stock returns. In addition, they also find that analysts tend to watch more close to those stocks with better risk-adjusted returns and usually revise their earnings forecasts more frequently for these stocks. These results imply that risk-adjusted returns of stocks have a significant influence on analysts' stock selection process.

2.2 Stock Price Informativeness

Richard G. Sloan (1996) examine whether stock prices present about future earnings within the accrual and cash flow components of current earnings. The results show that the relative magnitudes of the cash and accrual components of current earnings exert an impact on the extent to which current earnings performance persists in to the future. However, further evidence indicate that stock prices can't present fully informativeness involved in the accrual and cash flow components of current earnings until the information affects future earnings.

Joseph D. Piotroski (2000) asks the question that can a relatively simple accounting-based fundamental analysis strategy applying to a broad portfolio of high book-to-market firms, change the distribution of returns earned by an investor? The results of the study indicate that selecting financially strong high book-to-market firms could benefit investors by increasing the mean return by at least 7.5% annually, while the distribution of actual returns is shifted to the right. Moreover, evidence show that an investment strategy of longing expected winners and shorting expected losers generate a 23% annual return over the twenty years from 1976 to 1996, robust across time.

2.3 Links Between Analyst Coverage and Future Stock Returns

Lots of researchers have conducted study on the association between analyst coverage and future stock returns both in developed markets and emerging markets, however, with different conclusions due to many factors, such as different level of independence of the markets, studying on different periods, applying different models, etc. Most of the studies show relation between analyst coverage and future stock returns, some of which show positive relation while others show negative relation. A few studies show neutral relation between analyst coverage and future stock returns.

2.3.1 Global Stock Market

Positive relation:

The earliest study on the relation between analyst behavior and stock return is from James. H. B Jerring (1983). The subject of their study is the stocks analysts recommended. The results show that the information of recommendation does not reflect in the market price immediately. However, security investors investing in recommended stocks could gain statistically significant positive abnormal profit after deducting transaction costs. The limitations of this study is that the recommended stocks are mainly focusing on fossil energy industry, forestry and mining industry, which obviously could not represent the stock market as a whole.

Kent L. Womack (1996) focus on analysts in 14 major security companies and studies on their recommended stocks both with "buy" rating and "sell" rating. The results show significant relation between analysts' recommendation and both short-term future price and short-term trading volume. In addition, there is obvious systematical difference between price before recommendation and long-term value. For buy recommendations, the mean post-event drift is modest (+2.4%) and short-lived, but for sell recommendations, the drift is larger (-9.1%) and extends for six months. Analysts appear to have market timing and stock picking abilities.

Narasimhan Jegadeesh, et al (2001) suggest that sell-side financial analysts tend to recommend "glamour" stocks with positive momentum, high growth rate, high trading volume, and relatively high price. The extent of the consensus recommendation adds value only among stocks with positive quantitative characteristics such as high value and positive momentum. For stocks with negative quantitative characteristics, higher consensus recommendations are generally followed by poor performance in return. The quarterly change in the consensus recommendation seems to be a robust predictor of return which contains information orthogonal to a large range of other predictive factors.

Brad Barber, Reuven Lehavy (2002) study the possibility to benefit from investing in portfolios built by security analysts. They find that longing those high-rating recommended stocks, shorting those low-rating stocks and adjusting position every day in accordance with the changing recommendation portfolio can yield 4% annual abnormal profit. Decreasing in the frequency of position adjustment as well as a delay in reacting to recommendation changes eliminates the abnormal profit. Meanwhile, this trading strategy could generate high transaction costs, the portfolio profit after deducting the transaction costs is not statistically significantly greater than zero.

Karl B. Diether, Christopher J. Malloy (2002) suggest that stocks with lower dispersion in analysts' earnings forecasts earn higher future returns than that of stocks with higher dispersion. This conclusion is mostly found in small stocks and stocks that have performed poorly over 2001. They interpret dispersion in analysts' forecasts representing for differences in views about a stock. The study also suggests that this evidence is in accordance with the assumption that prices will reflect the optimistic view whenever investors with the lowest valuations trade or not.

Kalok Chan and Allaudeen Hameed (2004) also find that when the forecast dispersion is high, analyst coverage exerts less impact on stock price synchronicity. They investigate the association between the stock price synchronicity and analyst behavior in emerging markets. In contrast to the previous study that security analysts specialize in exploring firm-specific information, their results suggest that securities which are covered by more analysts incorporate greater (lesser) market-wide (firm-specific) information. They use the R-square statistics of the market model to measure the synchronicity of stock price movements, and the results suggest that more analyst coverage can lead to an increase in stock price synchronicity.

In addition, they also find that the aggregate changes in the earnings forecast of the high analyst-following portfolio exert a significant impact on the aggregate returns of the portfolio itself as well as those of the low analyst-following portfolio, while no predictability is showed from the aggregate changes in the earnings forecasts of the low analyst-following portfolio.

Randolph B. Cohen, Paul A. Gompers (2002) suggest that institutions as a group outperform individuals by only 1.44% per annum before transaction costs although institutions are trading in the ''right'' direction. Investigating the joint behavior of returns, cash-flow news, together with trading between individuals and institutions, they find that institutions buy shares from individuals in accordance with positive cash-flow news and sell shares to individuals in response to negative cash-flow news. Further evidence show that institutions are not simply following price momentum strategies. Institutions also sell shares to individuals when price goes up though in the absence of any cash-flow news, and vice versa.

David Easley and Maureen O'Hara (2004) show that by choosing features like accounting treatments, analyst coverage, and market microstructure, firms can exert an influence on their cost of capital. They examine whether and how information show influence on a company's cost of capital. The results show that the cost of capital could be affected by the differences in the composition of information between public and private information, as investors demanding a higher return to hold stocks with greater private information. The explanation for the fact that this higher return arises is that those investors who have access to be informed are much easier to modify their portfolio to incorporate new information, whereas uninformed investors are in the disadvantaged situation. In equilibrium, the quantity and quality of information affect asset prices.

Somnath Das (2006) investigates the predictability of financial analysts on future firm performance, based on the selective coverage of newly public firms. One key hypothesis that he makes is that analysts make their decision to provide coverage by mainly concerning their true underlying expectation of the future prospects of firms. They firstly get residual analyst coverage from a model of initial analyst following for newly listed firms, and then extract this underlying expectation, which is unobserved. The results suggest that in the subsequent three years after initial coverage, IPOs with higher residual coverage have significantly better return and operating performance than those with lower residual coverage.

Anna Scherbina (2007) presents evidence of inefficient information processing in equity markets by documenting that negative information withheld by securities analysts is incorporated in stock prices with a significant delay. She estimates the extent of the withheld negative information based on the proportion of analysts who stop revising their annual earnings forecasts. This measure predicts negative earnings surprises and negative price reaction around earnings announcements. It could also be used to generate profitable trading strategies. It shows that institutions tend to sell their stock holdings as my measure of unreported negative news increases, thus ameliorating the mispricing.

George Serafeim and Christopher Small (2010) present evidence that for stocks with lower price elasticity: i) return-revision synchronicity is lower, suggesting that analysts understand the noise in returns, ii) a trading strategy that uses forecast revisions delivers higher abnormal returns, and iii) analysts with high return-revision synchronicity make less accurate forecasts relative to analysts with high return-revision synchronicity for stocks with higher price elasticity.

Nont Dhiensiri (2010) finds that the price impact around a coverage initiation is positively related to the change in liquidity. Unlike the coverage initiations around the initial public offers (IPOs), the price impact is not related to the reputation of the analyst firm, the exchange listing or whether the analyst firm is also the IPO underwriter. The sample firms do not experience significant reduction in the level of information asymmetry but experience a significant increase in liquidity. The increase in liquidity only occurs after the coverage initiations. The increase in liquidity is not explained by the increase in institutional investors' interest.

Michael J. Jung (2015) shows that analyst interest is a novel and early indicator of future firm fundamentals and capital market consequences. He measures increases in analyst interest by observing analysts who do not cover a firm but participate in that firm's earnings conference call, and measures decreases in analyst interest by observing analysts who cover a firm, yet are absent from that firm's call, and finds that increases in analyst interest are positively associated with future changes in firm fundamentals and capital market activities, while decreases in analyst interest are negatively associated with capital market activities. He also finds that increases (decreases) in analyst interest are positively (negatively) correlated with future stock returns over the next three months and that a hedge portfolio yields a significant abnormal return.

S.P. Kothari (2016) indicates that 1) analysts' forecasts show predictable biases; 2) the market appears to underreact to the information in forecasts and to not fully filter the biases in forecasts; 3) Analysts' forecasts show positive effect in estimating expected returns on stocks, however further study on the association between analysts' forecasts and expected returns is still needed.

Charles M.C. Lee and Eric C. So (2016) suggest that analyst coverage proxies contain information about future expected returns. Decomposing analyst coverage into observed expected coverage and unobserved abnormal coverage applying a characteristic-based model, their results show that firms with abnormally high analyst coverage outperform firms with abnormally low analyst coverage by around 80 basis points per month for the subsequent months. Further evidence suggest that abnormal coverage rises following exogenous shocks to underpricing and predicts improvements in firms' fundamental performance, indicating that return predictability stems from analysts more heavily covering underpriced stocks. These conclusions prove the usefulness of analysts' actions in expected return estimations.

Francisco Marcet (2017) finds that analyst coverage networks (ACN) play an important role in explaining stock return commonalities across Latin American stocks. He finds that 1) higher co-movement is shown when analysts connect pairs of stocks; 2) Those stocks that foreign investors have easy access to trade appear to be affected heavily by common coverage; 3) international analysts are an important source of cross-country excess co-movement. By creating the network at the brokerage house level and exploiting exogenous changes in ACN around the MSCI LATAM Index reviews, the study further highlights endogeneity concerns related to the effect of ACN on commonalities.

Negative relation:

Harrison Hong et al (2000) find three main results based on the gradual-information-diffusion (GID) model of Hong and Stein (1997). 1) The return of momentum strategies declines sharply with firm size when one moves past the very smallest stocks, where thin market-making capacity could be an issue. 2) Controlling size fixed, momentum strategies exhibit a well performance among stocks with low analyst coverage. 3) A strong asymmetry is showed that the effect of analyst coverage usually works for stocks that performed poorly in the past rather than for stocks that performed well.

Ambrus Kecskés and Kent L. Womack (2006) show that the number of analysts following is negatively associated with future stock returns. Evidence shows that the decrease-increase return spread is 6.4 percentage points. The overreaction appears to be most pronounced when changes in analyst following are proved by changes in analysts' consensus recommendations or changes in institutional ownership. In addition, the overreaction also depends on valuation levels.

Neutral relation:

R Michaely and KL Womack (1999) find that stocks that underwriter analysts recommend perform more poorly than "buy" recommendations by unaffiliated brokers prior to, at the time of, and subsequent to the recommendation date. They believe that the recommendations by underwriter analysts show significant evidence of bias and market does not recognize the full extent of this bias. This finding indicates a potential conflict of interest in the different functions that investment bankers perform.

Ahmed Marhfor et al (2013) find that analysts' activities do not contribute to the impounding of future earnings information into current stock prices. They examine whether more analyst coverage translates into more informative stock prices and apply this to both developed and emerging markets. They measure price informativeness using the association between current stock returns and future earnings and argue that more informative stock prices contain more information about future earnings. The results is in accordance with the view that analysts are outsiders who do not have full access to firm-level information, and also consistent with the explanation that analysts focus on gathering and mapping industry- and market-level information (macroeconomic information) into stock prices.

Oya Altınkılıç (2016) finds that analysts tend to modify their forecasts of future long-term returns on their recommendations in the opposite direction. He investigates post-revision return drift, known as PRD. PRD refers to that the analysts change forecasts of future long-term returns in the same direction as the change in actual future return, to be specific, upgrading the rating are usually followed by positive returns, and downgrading the rating are generally followed by negative returns. He finds that during the high-frequency algorithmic trading period of 2003–2010, average PRD is no longer significantly different from zero.

2.3.2 China's Stock Market

Zhenshan Wang and Qiu Yao (2008) examine how analyst coverage affects stock return in Chinese A share. They believe that analyst coverage could provide more and accurate information compared with performance forecast and investing ratings. Analyst are likely to keep close watch on those firms with large size and good performance. Applying residual analyst coverage model, they find that stocks with higher analyst coverage have statistically significant positive higher return over those stocks with lower analyst coverage. The zero-investment portfolio they build based on this result could provide significant positive return. They also find that residual analyst coverage has positive correlation with stock return by applying Fama Macbeth regression. However, further research on the influence path of the two variables is still needed.

Rong Ding et al (2013) investigate the association between analyst coverage and stock price informativeness in China. The analysis of a sample of Chinese listed firms between 2003 and 2008 supports their conjecture that there is a positive association between stock price informativeness and analyst coverage. They further find that such association is more pronounced in non-state-owned enterprises (NSOE i.e. private firms) in that NSOEs are more dependent on external equity capital for financing and therefore have to maintain good relationship with analysts by timely responding to their enquiry and request, which results in more firm-specific information being incorporated into stock prices. They also show that the association between analyst coverage and stock price informativeness is more pronounced in less developed regions, which indicates that analyst coverage plays a more significant role in enhancing corporate information environment in regions where investor protection is weak.

Nianhang Xu et al (2013) examine the relations among analyst coverage, analyst optimism, and firm-specific stock price crash risk. Using a unique Chinese database, they find that an increase in a firm's analyst coverage leads to an increase in stock price crash risk and this positive relation is more pronounced when analysts are more optimistic analysts and are affiliated with investment banks and brokerage firms with mutual funds relation. They also find some weak evidence to suggest that analyst optimism on crash risk is less pronounced when analysts have high personal reputations or are affiliated with reputable brokerage firms.

Xunan Feng, Na Hu and Anders C. Johansson (2015) further study on the association between the effect of analyst coverage and ownership of firms and the results show that the extent of separation of control and ownership rights is positively related with the response coefficient of stock return synchronicity to analyst coverage.

Xuelian Bai et al (2016) prove the effectiveness of analysts’ role as producers of firm-specific information in Chinese IPO market and find that this role depends on the institutional environment. They select the data set from 2005 to 2012 covering the year 2009 where an important IPO regulation changes in China, the results show a significantly different effect of analyst coverage on synchronicity before and after the implementation of the IPO regulation. In particular, they find that analyst coverage decreases synchronicity with this effect significant only after 2009. Moreover, they further distinguish the information production role of underwriter and independent analysts and find that prior to 2009, underwriter analysts’ coverage reduces synchronicity while independent analysts’ coverage does not show impact on synchronicity. However, after 2009, both types of analyst coverage are significantly associated with synchronicity.

Mingshan Zhou et al (2016) examine star analyst coverage, investor overreaction, and stock price synchronicity in the Chinese and US markets. In China, they find that star analyst coverage can induce investor overreaction, such that it is negatively correlated with price synchronicity. This overreaction effect is particularly pronounced for stocks with primarily individual investors. In contrast, in the United States, they find that star analyst coverage is positively related to synchronicity and is not significantly associated with investor overreaction. The overall findings imply that the heterogeneous nature of investors in a market drives the association among star analyst coverage, overreaction, and stock price synchronicity.

3 Methodology and Data

In this paper we want to examine the link between analyst coverage and future stock returns. In this section, we first briefly introduce our primary methods/approaches/models applied to explore and test the link, at the same time discuss the required inputs of those methods/approaches/models. Then we talk about how we collect and clean input data prepared for use.

3.1 Methodology

Five methodologies mainly used in our study are introduced and discussed respectively in the following parts below.

3.1.1 Proxies for Analyst Coverage

Previous studies provided many measures of analyst coverage (Hong et al. 2000; Somnath et al. 2006; Wang and So, 2008; Lee and So, 2016). The simplest one is just counting and recording the number of unique analyst following a firm within a time interval (Somnath et al. 2006). An alternative way to proxy for analyst coverage is to count the total number of forecasts of a firm given by analysts within a time period (Wang and Yao, 2008. The data of the number of forecasts was gathered from different sources and manually counted[7].). The prior two measures are quite simple but unclear about how they deal with analysts' revisions. Lee and So (2016) applied a more complicated proxy for analyst coverage by measuring analyst coverage as the number of unique earnings forecasts summed across all analysts and forecasted fiscal periods (in analyst/forecast pairs, where revisions are single counted), incorporating in the proxy the extent to which analysts devote greater resources by forecasting earnings for more fiscal periods. But they also took into account the simple measure of analyst coverage used by Somnath et al. and found the results were qualitatively similar to those when using the more complicated measure. A similar concept to analyst coverage is analyst interest. Sometimes they are the same and exchangeable. But analysts can be interested in a firm but choose not to cover that firm. Jung (2015) came up with a measure for analyst interest. Increase in analyst interest is measured by the number of analysts who do not cover a firm but participate in that firm's earnings conference call. Decrease in analyst interest is measured by the number of analysts who cover a firm but are absent from the conference call.

[7] This time-consuming and energy-consuming way of data collection is the reason why they had only one-year analyst coverage data in their sample.

I/B/E/S Detail History offers detailed historical forecasts data and serves as the main data source in most previous studies about developed markets (Hong et al. 2000; Somnath et al. 2006; Lee and So, 2016). For emerging market, I/B/E/S International provides data on analyst activities for companies around the world and could be a good data source (Hameed et al. 2004). But after we carefully checked the data on Chinese listed companies from I/B/E/S International, we found some companies were missing in that database. So we decided to turn to some Chinese local database for more complete data. The comprehensive database we have access to is Wind[8]. It's quite similar to Bloomberg but focus on Chinese stock market. The problem with Wind is that it only provides data on the number of analysts (institutions) covering a firm and no more details. Due to this limitation, we have no choice but use the simplest proxy for analyst coverage, that is the number of analysts covering a firm within a time period and give up more complicated measures.

3.1.2 Abnormal/Residual Analyst Coverage Model

Total analyst coverage can be decomposed into two parts (Hong et al. 2000; Somnath et al. 2006; Wang and Yao, 2008; Lee and So, 2016). One is the part driven by commonly known determinants, most of them related to firm-specific characteristics, such as firm size, past performance, stock liquidity, etc. (Bhushan, 1989; Alford and Berger, 1999; Hong and Stein, 2000; Lee et al., 2001; Hameed et al. 2004; Somnath et al. 2006; Zhu et al., 2016; Lee and So, 2016; Rahadi and Lo, 2017). The other is the remaining portion of analyst coverage, unexplained by aforementioned determinants, but driven by other unknown or unproved factors. (Hong et al. 2000; Somnath et al. 2006; Wang and Yao, 2008; Lee and So, 2016). For example, self-selective analyst coverage, first formally documented by O'Brien and Macnichols in 1997, further discussed by Somnath et al. in 2006 and Lee et al. in 2016, leads analysts to distribute abnormally high coverage to firms for which they have favorable expectations but abnormally low and even no coverage to firms with unfavorable expectations. That remaining portion refers to *residual analyst coverage* in Somnath et al (2006) and *abnormal analyst coverage* in the paper of Lee and So (2016). The model used to decompose total analyst coverage and extract these two parts respectively is the abnormal/residual analyst coverage model.

The abnormal/residual analyst coverage model was first presented in the paper of Somnath et al. (2006). They developed this model inspired by the paper of Hong et al. in 2000,

---

[8] See http://www.wind.com.cn/en/Default.html for more information about Wind Financial Terminal and Database. And we will introduce Wind in more details in 3.2.1 Data Source.

where Hong regressed analyst coverage on firm size to obtain residuals. The abnormal/residual analyst coverage model regresses total analyst coverage on those commonly known determinants of analyst coverage. The model-based expected analyst coverage is used as the proxy for the portion of total analyst coverage determined by known factors. The residual from the model controlling those factors is used to proxy for abnormal analyst coverage, the portion not attributable to known determinants but potentially driven by other unknown or unproved factors. Because Somnath et al. (2006) mainly study the link between analyst coverage and future stock returns in the 7^{th} -24^{th} month subsequent to the IPO offering, the residual analyst coverage model presented by Somnath et al. includes many IPO-specific firm characteristics as explanatory factors in the model. And thus their model, though useful when analyzing analyst coverage in new issue market, is not widely applicable in other situations. Lee and So (2016) built upon Somnath's model but further improved and adjusted the residual analyst coverage model to be broadly applicable in cross-sectional tests to a greater number of firms, including firms with zero analyst coverage, in developed stock market. The three factors they included in the model are firm size, trading turnover, and firm's cumulative market-adjusted return (proxy for past performance). Their specified model is shown below.

$$Log(1 + TOT_{i,m}) = \beta_0 + \beta_1 SIZE_{i,m} + \beta_2 TO_{i,m} + \beta_3 MOMEN_{i,m} + \varepsilon_{i,m}$$

where TOT represents total analyst coverage of the firm i in month m, SIZE is the log of market value, TO is share turnover, and MOMEN is the cumulative market-adjusted return over the past 12 months. The residual term $\varepsilon_{i,m}$ is the abnormal analyst coverage.

We hypothesize that analyst coverage contains information about future stock returns and can be a robust return predictor. As we decompose analyst coverage into two parts, we can not only test the relationship of total analyst coverage with future stock returns, but also further examine whether the information about future stock returns comes from the unobservable, unexplained, abnormal analyst coverage. Somnath et al. (2006) document a significant relation between abnormal analyst coverage and future returns of newly public firms. Lee and So (2016) report that in the US market abnormal analyst coverage is significantly correlated with future firm performance and future stock returns, whereas total analyst coverage is not.

We focus on Shanghai Exchange A-share market of China, instead of the US market (Lee and So, 2016) and the new issue market (Somnath et al. 2006), and hence the determinants included in our residual/abnormal analyst coverage model should be different from those used by them. We will discuss in more details in the next part about the factors we should put in the residual/abnormal analyst coverage model.

3.1.3 Firm Characteristic Factors that Affect Analyst Coverage

Existing studies have shown many determinants of the total number of analysts following a firm (the simplest measure of analyst coverage). Bhushan (1989) find that analyst coverage level is an increasing function of firm size, the beta value of the stock, return volatility. Alford and Berger (1999) discuss that high trading volume of a stock, as a proxy for brokerage commission, might trigger more analysts to collect and supply information about the stock. Lee et al. (2001) report that analysts generally prefer 'glamour' stocks, those with positive momentum, high trading volume measured by turnover ratio, high ROA or ROE, and high value multiples such as book to market ratio and P/E. Lee and So (2016) examined many determinants mentioned above and show that, after including firm size, share turnover and firm's cumulative market-adjusted return in their residual analyst coverage model, adding any other firm characteristic factors such as return volatility, book to market ratio, ROA and alike gave little significant incremental explanatory power (measured by the value of R-squared of the model) for variation in total analyst coverage. So for parsimony they omitted other factors except size, share turnover and firm's cumulative market-adjusted return.

We re-examine many firm characteristic factors that are known as the explanatory variables for total analyst coverage (denoted as TCOV hereafter), the dependent variable of the residual analyst coverage model, and check whether iteratively including one of these factors significantly increases the explanatory power of our residual analyst coverage model (See Figure VIII in Appendix I). Finally, the four factors we selected to include in our residual analyst coverage model are firm size, share turnover, cumulative market-adjusted return, and ROA. So our residual analyst coverage model is specified below.

$$Log(1 + TCOV_{i,m}) = \beta_0 + \beta_1 SIZE_{i,m} + \beta_2 TO_{i,m} + \beta_3 MOMEN_{i,m} + \beta_4 \mathrm{ROA}_{i,q} + \varepsilon_{i,m}$$

where $TCOV_{i,m}$ represents total analyst coverage of the firm i in month m, $SIZE_{i,m}$ is the log of market value, $TO_{i,m}$ is share turnover , and $MOMEN_{i,m}$ is the cumulative market-adjusted return over the past 12 months in month m. The residual term $\varepsilon_{i,m}$ is the abnormal analyst coverage we are interested in, denoted as ACOV hereafter.

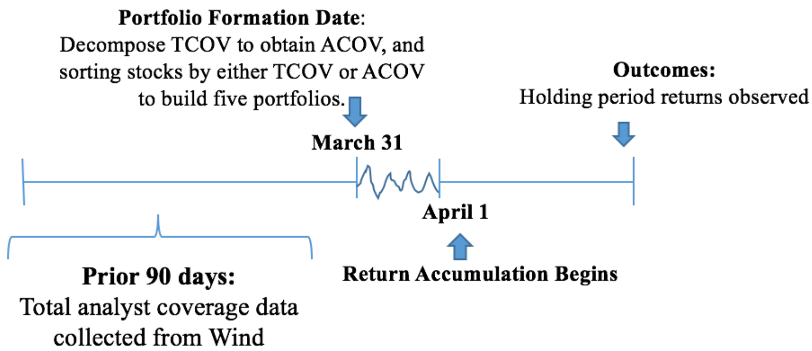
The left-hand side variable is *Log* (1 + $TCOV_{i,m}$). We take logarithm here because we assume that one extra analyst should matter more for a firm that has few analysts covering it than for one that is followed by a number of analysts (Hong, Stein. 2000). $SIZE_{i,m}$ is the value of market capitalization after logarithm. Using a logarithmic scale of firms' market capitalization here reduce the range of this variable's values when other variables in the model have relatively small scales.

We decompose total analyst coverage (TCOV) and derive abnormal analyst coverage, $ACOV_{i,m}$ from this model. If analysts do have the superior ability to pick firms with promising prospects and do cautiously select the firm they cover (Analysts' selective coverage), high (low) abnormal analyst coverage (ACOV) should be followed by superior (inferior) stock performance. This constitutes another key hypothesis: analyst coverage, especially ACOV, should be positively correlated with future stock returns.

3.1.4 Portfolio Sorts vs. Fama-Macbeth Panel Regression

We examined the link between analyst coverage and future stock returns through *Portfolio Sorts* and *Fama-Macbeth Panel Regression*. These two methods are commonly used to analyzing return anomalies (Fama and Macbeth. 1973, Fama and French. 2008). Portfolio sorts, sorting stocks by an anomaly variable and then calculating average returns, shows a picture of how average returns vary across the spectrum of an anomaly variable. Fama-Macbeth Panel Regression is an alternative to Portfolio sorts. The Fama-Macbeth regression applied here estimates parameters in two steps: 1) running cross-sectional regressions every period on time-variate firm characteristic variables such as book to market ratio, abnormal analyst coverage, etc. 2) calculating time-series averages to aggregate over time. One advantage of regression is that regression slopes can directly present the estimated marginal effects of anomaly variables on returns. These two methods can serve as a cross-check for each other.

In our study, we built five portfolio sorted by abnormal analyst coverage every month from Jan 2008 to Nov 2017 and calculated the equal-weighted average one-month-ahead returns for each portfolio to detect whether there is any pattern in return variation across abnormal analyst coverage quintiles. A hedge portfolio was also built. Below the diagram shows the timeline of our portfolio sorts analysis.



As mentioned in 3.1.1, we measure total analyst coverage as the total number of analysts (institutions) covering a firm over a period. To be more specific, we count analyst coverage over a 90-day window (about a quarter) at the end of month m. This choice of time window is to 1) allow reasonable time for analysts to digest information and to choose the firms they are willing to cover; 2) match the time interval between the release of last and new earning reports; 3) to match the time window high-quality databases (I/B/E/S International; *Wind*) and previous related studies (Wang and Yao, 2008; Lee and So, 2016) use when measuring analyst coverage. To avoid look ahead bias, we mimic what Lee and So did in their paper. Assuming we are at the end of month m now, i.e. March 31, we collect total analyst coverage data from *Wind* over prior 90 days. On March 31 we decompose total analyst coverage (TOCV) using the residual analyst coverage model to obtain abnormal analyst coverage (ACOV) and sort stocks based on either TCOV or ACOV. These five portfolios are held for a month before we replicate the process described above. By doing this, we ensure that all of the information used for portfolio sorts are observable and available prior to March 31 and all of the outcomes (returns) are observed after April 1.

We did Fama-Macbeth panel regression as a cross-check. We regressed one-month-ahead raw returns on total analyst coverage (TCOV) first and then with additionally controlling for firm size, share turnover, ROA, and cumulative market-adjusted returns over past 12 months. The regression with additional controls is equivalent to directly regressing returns on abnormal analyst coverage (ACOV). We expect that the coefficients of TCOV and ACOV are significantly different from zero. Our hypothesis can be confirmed by either TCOV significantly correlated with one-month-ahead raw returns or ACOV's significant relationship with future returns. And we further hypothesize that the coefficients are positive.

But we should also check whether the basic pattern found through portfolio sorts is affected by other dimensions such as size, book to market ratio, etc. In other words, we should test whether the predictive power of abnormal analyst coverage is distinct from other known predictors or determinants of stock returns. Fama four factors model serves this need.

3.1.5 Fama Four Factors Model

We adopt the classic Fama three factors model (Fama and French. 1993)with one more factor, the momentum factor, included (Carhart, 1997). We include momentum factor in our model based on the paper of Charles et al. (2001), where they report that the level of analyst recommendation derives its predictive power largely from a tilt towards high momentum stocks. Though our variable of interest is analyst coverage rather than analyst recommendation, we still think we should take momentum factor into account when examining the predictive power of analyst coverage for stock returns. The fama four factor model we adopt is:

$$Rpt - Rft = a + b_1 (R_{mt} - R_{ft}) + b_2 SMB_t + b_3 HML_t + b_4 \mathrm{MOM}_t + \varepsilon_t$$

where Rpt is the equally weighted returns of one of the five portfolios sorted by abnormal analyst coverage (or the returns of the hedge portfolio) in month t; Rft is the risk-free rate in month t. R_{mt} is the return of the A-share Index of Shanghai Exchange. *SMB, HML* and *MOM* are size factor, value factor and momentum factor at month t respectively from RESSET[9] database.

To test our hypothesis, we mainly examine whether the hedge portfolio earns a statistically significant alpha after controlling for the effect of other risk factors.

3.2 Data

In this part, we present our data source, data cleaning process and elaborate the data description.

3.2.1 Data Source

As mentioned before, most previous related studies turned to I/B/E/S Detail History for analyst coverage data, CRSP for stock returns and Kenneth R. French Data Library for Fama/French Factors data. But after we carefully checked the data on Chinese listed companies from I/B/E/S International, we found some companies were missing in that database. And we did not find the Fama/French factors data specifically for China’s A-share market of Shanghai Exchange in Kenneth R. French Data Library. So, we try China’s local database to gather more complete data. We mainly collect our data from two sources: Wind and RESSET. Wind, as the market

[9] More detailed introduction of RESSET in the next section 3.2.1 Data Source or see http://www.resset.cn .

leader in China's financial information service industry, provides accurate and real-time information for financial professionals. Its financial database offers the most comprehensive, complete data on Chinese stocks, bonds, funds, futures, RMB exchange rates, and the economy. We get the data we need for individual stocks, the data of monthly analyst coverage, monthly market capitalizations, monthly share turnover, etc. We use monthly data in our study because, compared with quarterly data or annual data, monthly data enlarged our sample size and give the tests in our study more statistical power. RESSET Database is a data platform provider of professional services for model test, investment research and so on. It's empirical research-oriented, with design idea, system structure, data quality, technical patterns reaching international advanced level. We use RESSET database to collect the data of Fama/French factors for China's A share market of Shanghai Exchange.

Table I provides an overview of the variables included in our study and their definitions.

Table I: Definitions and Computation Method of Key Variables[10]

	Definition	Computation Method Applied in This Paper
TCOV	Total Analyst Coverage, the total number of analysts who observe a particular stock during a period of time.	the number of institutions who have given a stock rating for a company in a given period. Specifically, we collect the total amount in 90 days prior to the specified days
SIZE	the total market value of a company's outstanding shares	closing price of A-share stock*number of A-share outstanding shares + closing price of B-share stock*number of B-share outstanding shares*RMB Exchange Rate + closing price of H-share stock*number of H-share outstanding shares*RMB

[10] Source of computation method: Wind

		Exchange Rate + closing price of stock listed in foreign stock market*number of all outstanding shares listed in foreign stock market*RMB Exchange Rate
TO	the trading frequency of an individual stock	the total number of shares traded over a given period (here we collect monthly data)/the average number of shares outstanding for the period
MOMEN	the difference between the return of a specified stock and the return of a specified market index over a given period	the return of stock over a given period – the return of a specified market index over that period (here a given period refers to a month, a specified market index refers to Shanghai Composite Index)
ROA	the amount of earnings returned as a percentage of total assets	EBIT*2/(opening balance of asset + closing balance of asset)*100%
ROE	the amount of net income returned as a percentage of shareholders' equity	Net Income (before dividends paid to common stock holders but after dividends to preferred stock)/shareholder's equity(not including preferred shares)
PE (TTM)	Price to Earnings Ratio (Trailing Twelve Months)	market price per share/earnings (over the prior 12 months) per share

PB	Price-to-Book Ratio	market price per share/((total Assets – total liabilities)/number of shares outstanding))
RETURN	the return over a given period including dividends reinvested	(closing price of the given period – opening price of that period)/opening price of that period*100% (here a given period refers to a month)

3.2.2 Data Cleaning

We collect 10-year monthly data of analyst coverage, returns and other firm characteristic variables for stocks listed in A-share market of Shanghai Stock Exchange (SSE).

We exclude from our sample the stocks that receive special treatment (denoted as ST or *ST stocks in SSE). Shares carrying "ST" tag suffer losses for two consecutive years or more. stocks with *ST are facing the risks of termination of listing. Chinese scholars usually treat those stocks as a special group distinct from other A shares (Wang and Yao, 2008). Given limited time, we only study normal A-shares. Further study can be done using specifically ST and *ST stocks as a sample.

We also drop stocks with frequent missing data during the sample period for a balanced panel data set. We are facing a trade-off – whether to go for a balanced panel by throwing away some pieces of usable information or to keep all usable observations in an unbalanced panel at the expense of methodological and computational complication (Hun Myoung Park, 2011). Both of us are not very familiar with unbalanced panel, so we go for a well-organized balanced panel data set, though we admit that dropping those stocks might be problematic. So by construction all stocks included in our sample survive throughout the sample period and thus don't have frequent missing data. But the cost is that we add some survivorship bias.

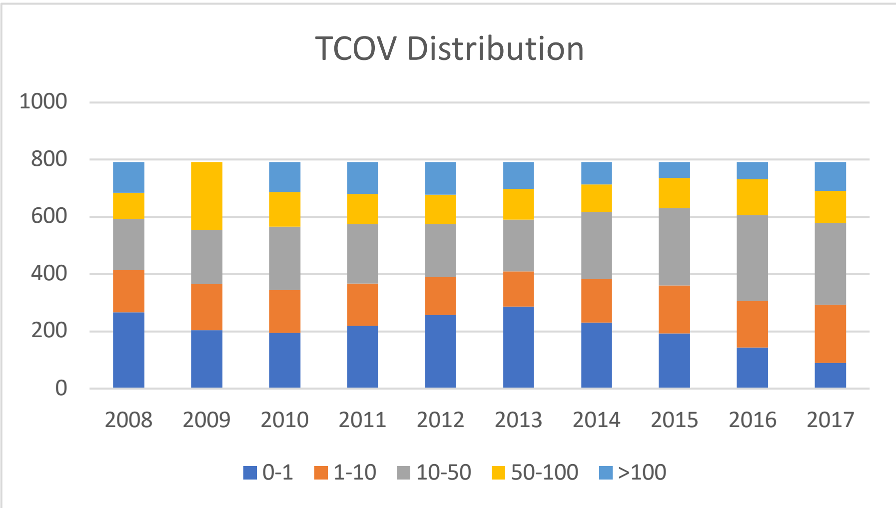
We keep those firms with zero analyst coverage, as Lee and So (2016) did in their research. For analyst coverage is the primary variable of interest, we think case-wise deletion based on analyst coverage will make us lose many information. And it's quite plausible that some public listed firms do not receive any coverage by analyst, especially when analysts' selective coverage exists.

Finally, we got 792 firms and their 10 year-monthly data as our sample, a 792 x 120 panel dataset.

3.2.3 Data Description

The total analyst coverage shows a relatively stable distribution over the past ten years[11]. Figure I below shows the yearly total analyst coverage from 2008 to 2017. (e.g., the yellow part in 2008 represents the number of firms whose total analyst coverage in 2008 is between 50 and 100.) We can see from the picture below that large total analyst coverage (>100) together with small total analyst coverage (0-1) are getting less, more total analyst coverage gather in the middle (1-100), which indicates that analyst coverage tend to be even across all the stocks listed in Shanghai A-share market.

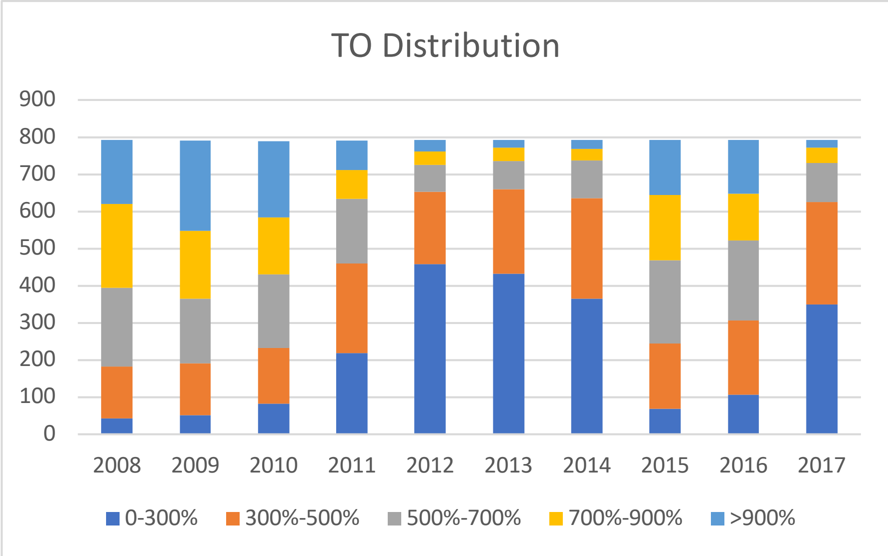
Figure I: TCOV Distribution



Share turnover has been through large fluctuations during the past ten years. Figure II below shows the monthly average share turnover over the past ten years. (e.g., the yellow area for 2008 represents the number of the firms whose monthly average share turnover in 2008 were between 50% and 100%.) From year 2008 to year 2012, share turnover has decreased at a stable rate, with average share turnover around 600%. Over the three years from year 2012 to year 2014, share turnover stays at a relatively lower rate around 300%. In 2015 and 2016, share turnover increased again. The average share turnover in these two years maintained around 600%. And then in 2017, share turnover dropped back to the level in 2014.

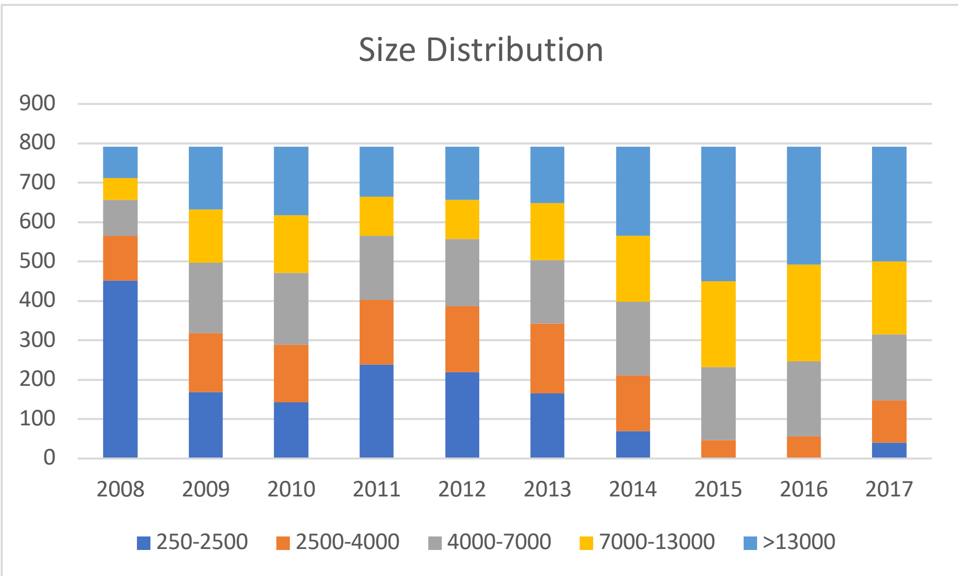
Figure II: TO Distribution

[11] See Appendix II for more details about the data



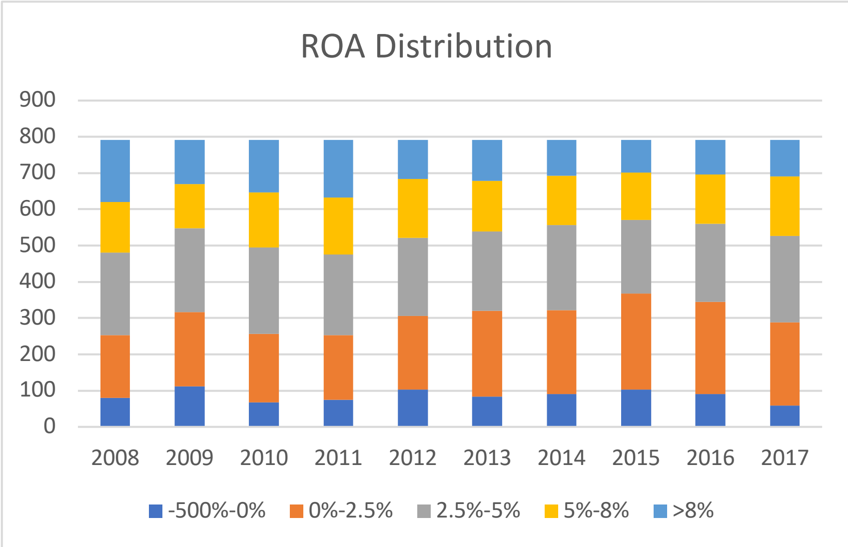
In general, the market capitalization of the firms listed in Shanghai A-share market are getting higher over the ten years. In 2008, the market cap over half of the firms are under 2500 million. The average market cap in 2008 was 14317 million. While after ten years, the average market cap increased by more than two times to 34641 million, which implicated a high-speed growth of China capital market from the perspective of market cap. Besides, we can see that Size distribution is positively associated with TO distribution.

Figure III: Size Distribution



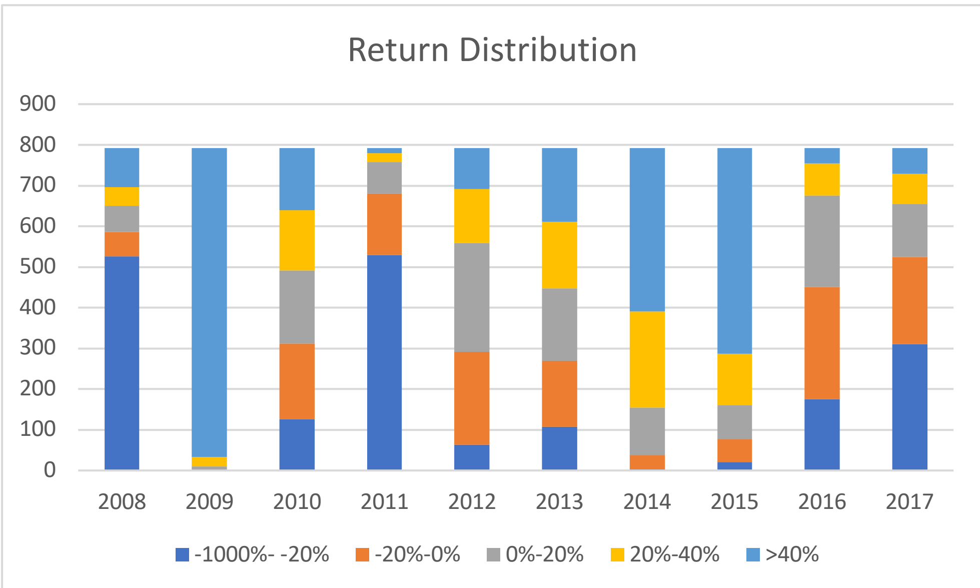
During the ten years, ROA stayed relatively stable. It shows that ROA did not fluctuate with neither share turnover nor size. The ten-year average ROA for the Shanghai A-share stocks was 4.1%.

Figure IV: ROA Distribution



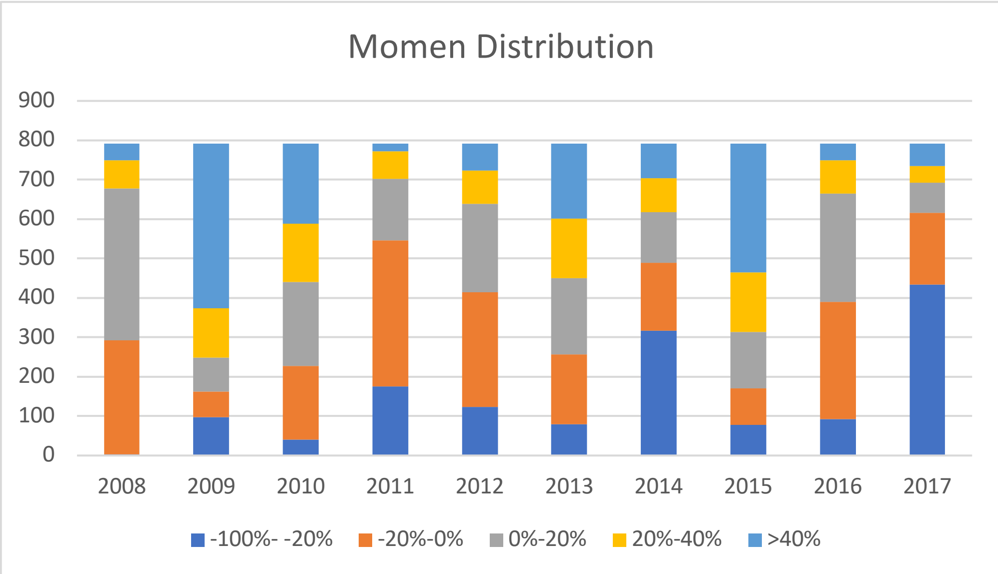
We can see rather large fluctuations in the return distribution over the ten years. Year 2008 was a disaster, while in the following year, which is year 2009, hardly did the firms end up with a negative return. Meanwhile, 96% of them had reached a return rate over 40%. In the year 2010, the return distribution is relatively of balance, with 61% of the firms got positive return. In 2011, the disaster happened again and the situation was even worse than that in year 2008. There was only 14% of the firms which got positive returns. From 2012 to 2015, stock returns over Shanghai A-share market kept increasing. In the following two years, the number of firms who got negative returns increased again. We can see that return distribution changes with an obscure pattern.

Figure V: Return Distribution



MOMEN distribution has the same moving pattern as that of Return distribution over the ten years.

Figure VI: MOMEN Distribution



Despite the detailed data description shown above, we also derived basic descriptive statistics of our sample from STATA. Table II presents basic descriptive statistics of the key variables in our study, where SIZE is measured by the logarithm value of market capitalization of firm i (i = 1, 2, 3, …, 792), RETURN are monthly returns at the end of calendar month m (m ranges from Jan 2008 to Dec 2012), ACOV is the proxy for abnormal analyst coverage derived from the residual analyst coverage model whose results are presented and discussed in 4.1.

Table II: Descriptive Statistics of Key Variables

	Obs.	Mean	Std. Dev	Min	Max
Log(1+TCOV)	95040	0.038	0.4170	0	1.5798
SIZE	95040	9.831	0.5244	8.3064	12.6335
TO	95040	5.369	3.7057	0	32.2372
MOMEN	95040	0.152	0.4904	-1.3503	12.0963
ROA	95040	0.034	0.0900	-9.0354	2.6779
RETURN	95040	0.0166	0.3431	-7.5719	55.5885

ACOV	95040	-0.0296	0.3385	-4.0994	13.7651

We have in total 95040 observations (a 792 firms * 120 months panel data set). It can be seen from Table II that there are some outliers/extreme values in ROA and RETURN. Some firms have negative ROA within the examination window of our study. Negative ROA could occur when a company is poorly utilizing and managing its assets. RETURN have a relatively low standard deviation compared with its range (Max-Min), indicating the existence of outliers. We come back to the raw dataset to check those outliers and find that they are densely located in the last month of 2008, a year that witnessed global financial crisis and a 4 trillion RMB stimulus plan announced by Chinese government. We also check with the database we use for data collection if any error or mistake occur when we export the data. Finally, we conclude that they are outliers in our sample. To deal with this problem, we formed a subsample excluding the Dec 2018 data for comparison, in order to see whether and how those outliers will affect our tests and results. We obtain qualitatively and quantitatively similar results using the sub-sample without data in Dec 2008 as the results using the complete sample from the residual analyst coverage model. But the results differ substantially when we sort stocks by analyst coverage with and without data in Dec 2008. We will discuss these further in the section of Empirical Results (4.2. and 4.3).

4 Empirical Results

Empirical results are presented in this section. The following parts report the results from our residual analyst coverage model, portfolio sorts, Fama-Macbeth panel regression and Fama four factor model respectively.

4.1 The Residual Analyst Coverage Model

Table III presents the time-series average coefficients from estimating the residual analyst coverage model shown in 3.1.3. All coefficients are significantly different from zero at 1% significance level. The regression results indicate that total analyst coverage[12] increases with firm size, past performance (MOMEN), and ROA. These positive correlations with total analyst coverage are consistent with the findings documented in previous studies (Bhushan.1989: the level of analyst coverage is positively associated with firm size; Charles et al. 2001: analyst prefer to cover stocks with positive momentum and high profitability.). By contrast, the coefficient of share turnover suggests that share turnover, a common proxy for stock liquidity, has negative impact on the level of total analyst coverage, which is in contrast with what Berger et al. 1999 and Charles et al. 2001 report in their papers that stocks with high trading volume/ high share turnover attract more analysts. Instead, this negative correlation indicates that analysts who are active in Shanghai Exchange A-share market prefer relatively illiquid stocks. Such coverage preference might attribute to the higher return earned by those relatively illiquid stocks (Amihud and Mendelson, 1991). Residuals from this model are taken as the proxy for abnormal analyst coverage (ACOV).

The average R-squared value reported in Table III shows that the residual analyst coverage model on average explains about 48% of the variation in total analyst coverage. In other words, expected analyst coverage account for around 48% of total analyst coverage, while abnormal analyst coverage constitutes the rest 52% of total analyst coverage, slightly over a half. This R-squared value is much lower than the one from the paper of Lee and So (2016) about the US market, which is over 60%. They include only three factors in their residual analyst coverage model: SIZE, TO and MOMEN. According to Lee and So, almost 58% of the variation of analyst coverage in the US stock market is explained by firm size, whereas our model shows that firm size only explains around 44% of the variation of analyst

[12] As mentioned in 3.1.2, we use log(1+TCOV) in the left-hand side of the equation of our residual analyst coverage model.

coverage in Shanghai Exchange A-share market. Though we include one more factor, ROA, in our model to explain variate analyst coverage, compared with the model adopted by Lee and So (2016), we still get a R-squared value below 50%. Analyst coverage in SH. A-share market of China appears to contain larger abnormal part that the counterpart of the US stock market.

Table III: Average Coefficients

	Mean	t-statistic
INT	-4.498	-62.74
SIZE	0.500	63.42
TO	-0.013	-25.76
MOMEN	0.062	6.89
ROA	1.482	14.13
R^2	0.4826	

We now successfully decompose monthly total analyst coverage (TCOV) and obtain monthly abnormal analyst coverage (ACOV) data: the residuals from the residual analyst coverage model. Basic descriptive statistics of ACOV can be seen in Table II in 3.2.3 and as mentioned in 3.2.3, the estimates of residual analyst coverage model when using the sub-sample without data of Dec 2008[13] are qualitatively and quantitatively similar to the estimates presented above.

4.2 Portfolio Sorts

To examine the link of analyst coverage (both TCOV and ACOV) with future stock returns, we move to next step: portfolio sorts based on analyst coverage.

4.2.1 Portfolio Sorts by Total Analyst Coverage

We rank stocks based on descending monthly total analyst coverage and then assign them to 3 groups building 3 portfolios where the "High" portfolio contains firms with high coverage and the "Low" portfolio consists of firms with low coverage. Because monthly abnormal analyst

[13] See Appendix III, Table IX

coverage data are derived from the residual analyst coverage model where the dependent variable is Log (1 + TCOV), to be consistent we sort firms based on total analyst coverage, Log (1 + TCOV), at the end of month m. Portfolio returns are calculated in month m+1. Time variable m ranges from Jan 2008 to Nov 2017, i.e. from 1 to 119. As showed by the timeline in 3.1.4, we mainly follow 1-0-1 portfolio construction rule, sorting stocks based on the past 1-month information, waiting 0 months before portfolio formation, and holding portfolios for another 1 month. Table IV presents equally weighted average monthly raw returns across 3 portfolios sorted by total analyst coverage, along with a hedge portfolio "High-Low" forming by longing the "High" portfolio at the same time shorting the "Low" portfolio. Corresponding t-statistics are also shown. We observe cross-sectional differences in one-month raw holding returns to detect if there is any pattern that suggests any expected return information embedded in analyst coverage.

Table IV: Equally Weighted Average Returns across Portfolios sorted by Total Analyst Coverage

Portfolios	1 (Low)	2	3 (High)	High-Low
EW Average Return	0.0217	0.0143	0.0111	-0.0106
T-statistics	1.7239	1.5201	1.3165	-1.6460

The results shown in Table IV suggest an insignificant but negative correlation between total analyst coverage and future stock returns. Stocks with lower analyst coverage seem to outperform those with higher analyst coverage on average. The "Low" portfolio has an average 2.17% return per month on an equal-weighted basis, the highest return among the three portfolios, and this mean return is statistically significant at 10% significance level with a t-statistics = 1.724. By contrast, the "High" portfolio averagely generates a monthly 1.11% return, the lowest among the three, and does not significantly different from zero. The average difference between the "High" portfolio and the "Low" portfolio is approximately 106 basis point per month on an equal-weighted basis, a difference both economically and statistically insignificant (t-statistic = -1.646). These findings indicate that total analyst coverage contains little information useful to predict future returns and the predictive power is insignificant. This implication is consistent with the finding of Lee and So (2016).

Results are qualitatively similar when we use the subsample without data in Dec 2008 instead of the complete sample (See Appendix III Table X). Total analyst coverage has little predictive power for future stock returns.

4.2.2 Portfolio Sorts by Abnormal Analyst Coverage

In this part, we examine the link between abnormal analyst coverage (ACOV) and future stock returns. We reassign stocks to quintiles of ACOV at the end of month m and calculate the cumulative return at the end of month m+1 (m ranges from 1 to 119, from Jan 2008 to Nov 2017). The higher quintiles correspond to stocks with relative high abnormal coverage while the lower quintiles correspond to stocks that receive abnormally low coverage from analysts. Table V contains equal-weighted average one-month ahead returns across abnormal coverage quintiles.

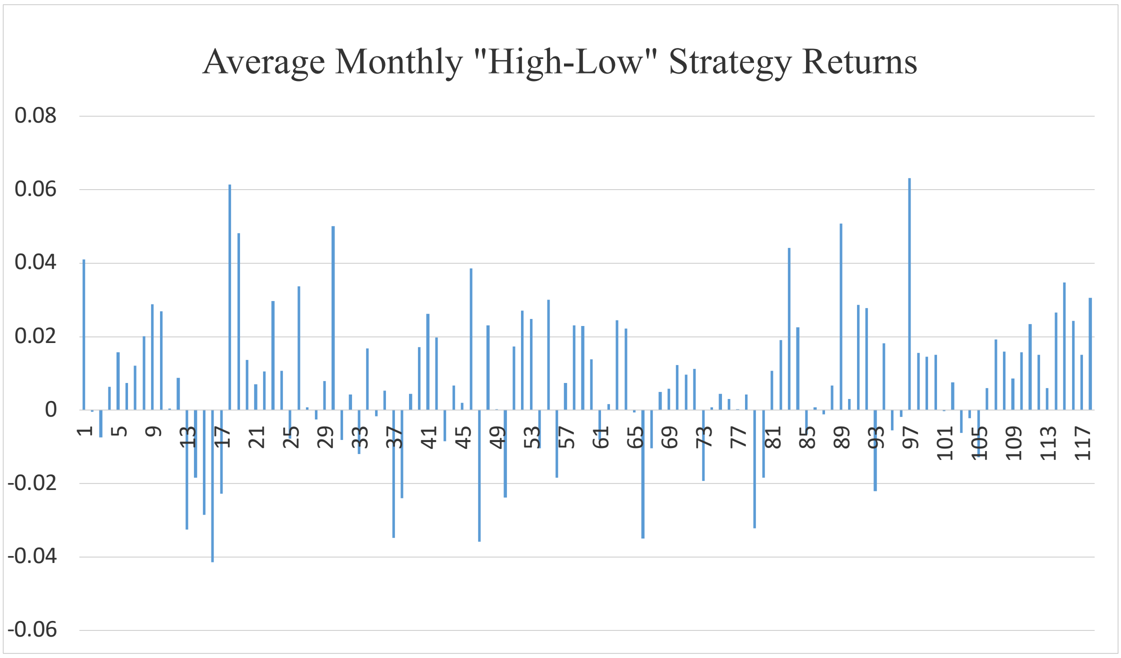
Table V: Equally Weighted Average Returns across Abnormal Analyst Coverage Quintiles

Quintiles/Portfolios	1 (Low)	2	3	4	5 (High)	High-Low
EW Average Return	0.0189	0.0153	0.01914	0.0167	0.0163	-0.0026
T-statistics	1.1710	1.4698	1.9008	1.8397	1.8427	-0.2396

No clear relation between ACOV and one-month-ahead returns can be observed from Panel B, with none quintiles generate average monthly returns significantly different from zero, given 5% significance level. And the average return difference between the highest quintile of ACOV and the lowest quintile is also statistically and economically insignificant. We plot the monthly returns from the strategy that long the highest quintile and short the lowest quintile, reflected by the “High-Low” column, to capture some meaningful pattern (See Appendix IV Figure XV). The strategy is implemented at the end of m and held in month m+1. we find that this strategy earned an extremely large negative return, 120% below zero, when held in Dec 2008, exactly the time point when many outliers/extreme values of returns occur, as we mention in 3.2.3. This non-recurring extreme return should not be take into account when we want to find stable pattern in the variation of a variable and test stable link between variables. If we treat the holding return in Dec 2008 as an outlier and rule out the effect from it, we have Table VI that reports equal-weighted average one-month ahead returns across abnormal coverage quintiles, and Figure VII that shows the monthly returns of the “High-Low” strategy.

Table VI: EW Average Returns across Abnormal Analyst Coverage Quintiles w/o the holding return in Dec 2008						
Quintiles/Portfolios	1 (Low)	2	3	4	5 (High)	High-Low
EW Average Return	0.0053	0.0102	0.0143	0.0139	0.0132	0.0079
T-statistics	0.6045	1.1130	1.6061	1.5972	1.5802	4.3040

Figure VII: Average Monthly "High-Low" Strategy Returns



Still, none quintiles generate significant average monthly returns at 5% significance level, but Table VI suggest a significant positive correlation between abnormal analyst coverage and future stock returns. Stocks in the highest quintile outperform those in the lowest quintile on average by around 79 basis points per month on an equal-weighted basis, with a t-statistic equal to 4.3040. Such a substantial difference corresponds to an annualized return of 10% if we implement the “High-Low” strategy. Such an annualized return is not only statistically significant but also economically significant. Evidence from Figure VII also support the positive link between abnormal analyst coverage and future stock returns. The one-month holding returns are generally positive. To be more specific, returns are positive in 83 out of the 118 months within our sample window. These results show that abnormal analyst coverage do contain information about future stock returns and could be a return predictor. These results also confirm our hypothesis that abnormal analyst coverage is positively correlated with future stock returns, suggesting that stocks that receive abnormally higher (lower) coverage from analyst are followed by higher (lower) returns. To some extent, this positive correlation indicates that analysts who are active in Shanghai A-share market do have some superior ability to pick stocks with promising prospect and do cautiously select the firms they decided to cover.

Sonmath et al. (2006) document a positive relation of abnormal analyst coverage with future stock returns in new issue market. Lee et al. (2016) also document this positive relation in developed markets. We now find evidence support this positive relation in Chinese market. But our findings are very sensitive to outliers/extreme values. Our results imply that this positive relation doesn't hold if we use the complete sample rather than the subsample without data in Dec 2008. In other words, the positive link between abnormal analyst coverage and future stock returns get broken when relatively rare events occur, e.g. 2008 financial crisis, while Lee and So (2016) argue that this positive link does not condition on rare events or specific context.

Results from portfolio sorts seems support our hypothesis if ruling out the effect of the extreme value in Dec 2008. We conduct Fama-Macbeth Panel Regression for cross-check.

4.3 Fama-Macbeth Panel Regression

As mentioned in 3.1.4, We conducted Fama-Macbeth panel regression, regressing one-month-ahead raw returns on total analyst coverage (log(1+TCOV)) first and then with additionally controlling for firm size, share turnover, ROA, and cumulative market-adjusted returns over past 12 months. The regression with additional controls is equivalent to directly regressing returns on abnormal analyst coverage (ACOV). Result are shown in Table VII.

Table VII: Time-series Average Coefficients from Fama-MacBeth Regression

	the Subsample Excluding Data on Dec 2008		*the Complete Sample*	
	(1)	(2)	(3)	(4)
Log(1+TCOV)	-0.0063	0.0052**	-0.0119	-0.0081
	(-1.70)	(2.31)	(-1.77)	(-0.60)
SIZE		-0.0186***		-0.0113
		(-5.08)		(-1.39)
TO		-0.0016***		-0.0016***
		(-4.22)		(-4.25)

MOMEN		-0.0111***		-0.009**
		(-3.48)		(-2.38)
ROA		0.0618***		0.0731***
		(2.76)		(2.94)
Intercept	0.0130	0.1996	0.0208	0.1407**
	(1.38)	(5.16)	(1.71)	(2.00)
R^2 (%)	1.88	6.85	1.87	6.81

The notation *** and ** indicate the coefficient is significant at the 1% and 5% respectively.

The results from Fama-Macbeth regression confirm our conclusion in 4.2. It can be seen from Column (1) and Column (3) that the relation of total analyst coverage, Log(1+TCOV), with future stock returns is insignificant in univariate tests. Results in Column (2) reports that the coefficient of total analyst coverage becomes significant and positive once we add SIZE, TO, MOMEN and ROA to the regression as control variables when the subsample excluding data in Dec 2008 is used, while column (4) shows an insignificant negative coefficient of total analyst coverage in multivariate tests when the complete sample is used in regression. Recall that the variation in total analyst coverage becomes equivalent to the variation in abnormal analyst coverage when we control for those firm characteristics. Again, we can conclude that, after excluding outliers/extreme values, unobserved abnormal analyst coverage contains information about future stock returns and is positively associated with one-month-ahead raw returns, whereas the correlation between observed total analyst coverage and future stock returns is insignificant and unclear.

In Column (2), coefficients of other variables can also be read. As these variables are included in the regression as control variables, we do not interpret their coefficients.

4.4 Fama Four Factors Model

We also need to examine the monthly strategy returns on an equally weighted basis we got in 4.2.2 controlling for the effect of known risk factors such as SMB, HML, etc. to see whether the return predictive power of absolute abnormal analyst coverage is robust or orthogonal to other asset pricing factors. We are interested in the alpha of the "High-Low" strategy longing the highest quintile of abnormal analyst coverage at the same time shorting the lowest quintile (re-balanced monthly), after controlling for returns of market portfolio, SMB, HML and MOM. The alpha corresponds to the intercept from the regression. The table below reports the results of Fama four factor regression analysis across time-series returns of each quintile of abnormal analyst coverage, showing factor-adjusted alpha, factor loadings on different asset-pricing factors and adjusted R-squared values for each model.

Table VIII: Fama Four Factor Regression Analysis

	a	RMRF	SMB	HML	MOM	R^2(%)
Lowest Quintile	-0.0052***	1.02***	0.70***	-0.19***	0.0074	96.77
	(-3.18)	(43.43)	(11.87)	(-2.92)	(0.18)	
Lower Quintile	-0.0023	1.02***	0.92***	-0.19***	-0.0037	96.62
	(-1.32)	(36.12)	(11.08)	(-2.52)	(-0.09)	
Medium Quintile	0.0016	1.00***	0.84***	-0.27***	-0.0488	96.87
	(0.97)	(43.65)	(12.99)	(-4.20)	(-1.08)	
Higher Quintile	0.0034**	1.02***	0.67***	-0.24***	-0.0079	96.52
	(1.98)	(43.54)	(8.53)	(-3.19)	(-0.18)	
Highest Quintile	0.0036*	0.99***	0.53***	-0.43***	-0.0080	95.26
	(1.90)	(34.65)	(6.94)	(-5.31)	(-0.16)	
High-Low Strategy	0.0089***	-0.031	-0.17**	-0.23***	-0.0154	14.56
	(5.23)	(-1.17)	(-2.53)	(-2.84)	(-0.34)	

The notation ***, ** and * indicate the coefficient is significant at the 1%, 5% and 10% respectively

Table VIII shows that one-month-head returns across abnormal analyst coverage quintiles have significant positive loadings on the market portfolio, close to 1, and significant positive loadings on size factor (SMB). But the loading on size factor decreases from the lowest quintiles to the highest quintile, indicating that the higher future returns of higher ACOV quintiles we observe in 4.2.2 partly stems from the fact that higher quintile of ACOV includes smaller firms. The loadings on value factor, HML, are statistically significant and negative, indicating that analyst active in Shanghai A-share market prefer to cover growth stocks.

Momentum factor appears to have little explanatory power to one-month-ahead returns across abnormal analyst coverage quintiles. By contrast, the zero-investment portfolio built based on the "High-Low" strategy has an insignificant loading, which is close to zero, on market portfolio, suggesting that this strategy does not track the broad market.

A monotonic increase in the factor-adjusted alpha across abnormal analyst coverage quintiles can be observed in the table above, growing from a negative value to a value above zero. Specifically, the factor-adjusted alpha of the zero-investment portfolio is highly significant different from zero and indicates that the average monthly return from the "High-Low" strategy remain both statistically and economically significant after we control for Fama four factors. The factor-adjusted excess return from such strategy is 89 basis points in an equal-weighted basis, slightly higher than the 79 basis points we obtain in 4.2.2 without any control factors. Such factor-adjusted alpha implies an annualized alpha of 11.22%.

In all, the predictive power of abnormal analyst coverage for future returns remain strong after controlling for other asset pricing factors and can be a return predictor robust to standard Fama four factors.

5 Conclusion and Discussion

In this paper, we move our eyes off analysts' activities in developed market such as the US stock market to focus on their activities in the Shanghai A-share market of China, a representative of fast-growing emerging markets. And instead of studying analysts' recommendations and forecasts about a firm (refer to 'what analysts say', in the paper of Lee and So in 2016), we choose analyst coverage decisions ('what analysts do', according to Lee and So, 2016) as the variable of interest. The self-selective coverage of analyst documented by O'Brien et al., as well as the theory on resources and attention allocation of security analysts discussed by Lee et al., inspire us to hypothesize that analyst coverage (Both total analyst coverage and abnormal analyst coverage) contains information (could be positive or negative signal) about future firm performance measured by future stock returns, as analysts cautiously select the firms they cover relying on their expectations about firms' future performance. Based on the positive correlation between residual/abnormal analyst coverage documented by Sonmath et al. in IPO market, Lee and So in developed markets, we further hypothesize that analyst coverage, especially abnormal analyst coverage, is positively correlated with future stock returns, suggesting that analysts do have sufficient skills to pick promising firms. We examine our hypotheses by implement 1) the broadly applicable residual analyst coverage model developed and improved by Hong et al. (2001), Sonmath et al. (2006), Wang and Yao (2008), Lee and So (2016), to decompose observed total analyst coverage into expected and abnormal parts; 2) Portfolio sorts and Fama-Macbeth panel regression to explore the relationship of future stock returns with total analyst coverage and abnormal analyst coverage respectively; 3) Fama four factor asset pricing model to test the robustness of the returns from strategy exploiting the link between analyst coverage and future share returns.

Main conclusions we draw from this study are 1) the correlation of total analyst coverage with future stock returns is neither statistically significant nor economically significant, while 2) abnormal analyst coverage, which is unobserved in the market, is positively associated with future stock returns after excluding outliers in Dec 2008, indicating that stocks that receive abnormally higher (lower) coverage from analyst are followed by higher (lower) returns, and the monthly-rebalanced strategy that longs stocks in the highest quintile of abnormal analyst coverage at the same time shorts stocks included in the lowest quintile earns an annualized return of approximately 10% on average, robust to standard asset pricing factors (RMRF, SMB, HML, MOM). Thus, abnormal analyst coverage can be a powerful return predictor. However, our results also show that this positive link doesn't hold if outliers in Dec 2008 are included in the sample, implying the positive link between abnormal analyst coverage and future stock returns is vulnerable to rare events, e.g. 2008 financial crisis, whereas Lee and So (2016) argue that this positive link does not condition on rare events or specific context.

Our research contributes to studies on analyst activities and behavior in emerging markets, especially Chinese stock market; and also supplements the empirical evidence of the positive link between abnormal analyst coverage and future stock returns, together with Sonamth et al. (2006), Wang and Yao (2008), Lee and So (2016). In addition, we provide a practical investment strategy to get advantage of this relationship. Moreover, the decomposition of total analyst coverage provides easily portable expected and abnormal parts of coverage for future studies. Still, further researches can be done on this specific topic towards these directions: 1) adding control for industry or do double sorts based on both industry and abnormal analyst coverage, because stocks in some promising, profitable industries might receive abnormally high coverage from analysts (Jerring, 1983), and the positive link between future stock returns and abnormal analyst coverage found in our thesis might be undermined once we control for industry, but we fail to take into account in our thesis the effect from that; 2) synergies between abnormal analyst coverage and analysts' recommendations, recall that both what analysts do and what analysts say are correlated with future returns (See 2. Literature Review); 3) the correlation of analyst coverage with future firm performance measured by other proxies or metrics other than future stock returns; 4) different proxies for analyst coverage, as total analyst coverage can be further categorized to coverage from different financial institutions and can be given different weights to constitute a totally new measure for total analyst coverage.

References

Somnath Das, Re-Jin Guo, and Huai Zhang, 2005, Analysts' Selective Coverage and Subsequent Performance of Newly Public Firms, The Journal of Finance 61(3), 1159-1185.

Huai-Chun Lo, Febri Rahadi, 2017, Analyst Coverage and Stock Returns, Journal of Business and Management Sciences. 2017, Vol. 5 No. 2, 42-56.

Michael J. Jung, M.H.Franco Wong, and X. Frank Zhang, 2014, Analyst Interest as an Early Indicator of Firm Fundamental Changes and Stock Returns, The Accounting Review, May 2015, Vol. 90, No. 3, 1049-1078.

S. P. Kothari, Eric So, and Rodrigo Verdi, 2016, Analysts' Forecasts and Asset Pricing: A Survey, Annual Review of Financial Economics, Vol. 8, 197-219.

Oya Altınkılıç, Robert S. Hansen, and Liyu Ye, 2016, Can analysts pick stocks for the long-run? Journal of Financial Economics, 119 (2016), 371-398.

Brad Barber, Reuven Lehavy, Maureen Mcnichols, and Brett Trueman, 2001, Can Investors Profit from the Prophets? Security Analyst Recommendations and Stock Returns, The Journal of Finance • VOL. LVI, NO. 2• April, 2001, 56-84.

Randolph B. Cohena, Paul A. Gompers, and Tuomo Vuolteenaho, 2002, Who Underreacts to Cash-Flow News? Evidence from Trading between Individuals and Institutions, Journal of Financial Economics 66 (2002), 409–462.

Lauren Cohen, and Dong Lou, 2012, Complicated Firms, Journal of Financial Economics, 104(2012), 383–400.

Karl B. Diether, Christopher J. Malloy, and Anna Scherbina, 2002, Differences of Opinion and the Cross Section of Stock Returns, The Journal of Finance, VOL. LVII, NO. 5 • OCTOBER 2002, 3-25.

Eugene F.Fama, and Kenneth R. French, Size, Value, and Momentum in International Stock Returns, Journal of Financial Economics, 105(2012), 457–472.

Ravi Bhushan, Firm Characteristics and Analyst Following, 1989, Journal of Accounting and Economics, Volume 11, Issues 2–3, July 1989, 255-274.

George Serafeim, and Christopher Small, 2010, Analyst Forecast Revisions and Noise in Stock Prices, The Accounting Review, 78(1) · April 2003, 3-24.

Harrison Hong, Terence Lim, and Jeremy C. Stein, 2000, Bad News Travels Slowly: Size, Analyst Coverage, and the Profitability of Momentum Strategies, The Journal of Finance, VOL. LV, NO.1, 354-381.

Mark T. Bradshaw, 2004, How Do Analysts Use Their Earnings Forecasts in Generating Stock Recommendations? The Accounting Review, Vol. 79, No. 1 (Jan., 2004), 25-50.

Malcolm Baker and Jeffrey Wurgler, 2006, Investor Sentiment and the Cross-Section of Stock Returns, The Journal of Finance, VOL. LXI, NO. 4 • AUGUST 2006, 135-172.

Narasimhan Jegadeesh, Joonghyuk Kim, Susan D. Krische, and Charles M. C. Lee, Analyzing the Analysts: When Do Recommendations Add Value? The Journal of Finance, Vol. 59, No. 3 (Jun., 2004), 1083-1124.

Roni Michaely, and Kent L. Womack, 1999, Conflict of Interest and the Credibility of Underwriter Analyst Recommendations, The Review of Financial Studies, Volume 12, Issue 4, 2 July 1999, 653–686.

Eugene F. Fama, and James D. MacBeth, 1973, Risk, Return, and Equilibrium: Empirical Tests, The Journal of Political Economy, Vol. 81, No. 3 (May - Jun., 1973), 607-636

Richard G. Sloan, 1996, Do Stock Prices Fully Reflect Information in Accruals and Cash Flows About Future Earnings? The Accounting Review, Vol. 71, No. 3. (Jul., 1996), 289-315.

Anna Scherbina, 2008, Suppressed Negative Information and Future Underperformance, Review of Finance (2008) 12, 533–565.

Mark T. Bradshaw, 2002, The Use of Target Prices to Justify Sell-Side Analysts' Stock Recommendations, Accounting Horizons, March 2002, Vol. 16, No. 1, 27-41.

Charles M.C. Lee, and Eric C. So, 2017, Uncovering Expected Returns: Information in Analyst Coverage Proxies, Journal of Financial Economics, Volume 124, Issue 2, May 2017, 331-348.

Kent L. Womack, 1996, Do Brokerage Analysts' Recommendations Have Investment Value? The Journal of Finance, Vol. 51, No. 1 (Mar., 1996), 137-167.

Zhenshan Wang, and Qiu Yao, 2008, Analyst Coverage and Stock Return, Research on Financial and Economic Issues, No. 4 (General Serial No. 293), 67-84.

Rong Ding, Wenxuan Hou, Jing-Ming Kuo, and Edward Lee, 2013, Does Analyst Coverage Affect Share Price Informativeness of Chinese Listed Firms? Regional Studies, 2017 forthcoming, 3-22.

Nianhang Xu, Xuanyu Jiang, Kam C. Chan, and Zhihong Yi, 2013, Analyst coverage, optimism, and stock price crash risk: Evidence from China, Pacific-Basin Finance Journal 25, 217–239.

Xunan Feng, Na Hu and Anders C. Johansson, 2015, Ownership, Analyst Coverage, and Stock Synchronicity in China, International Review of Financial Analysis, Volume 45, May 2016, 79-96.

Xuelian Bai, Fangjun Wang, and Junrui Zhang, 2016, Analyst coverage and stock return synchronicity: evidence from regulation changes in China's IPO market, Applied Economics, Volume 48, 2016 - Issue 47, 4538-4557.

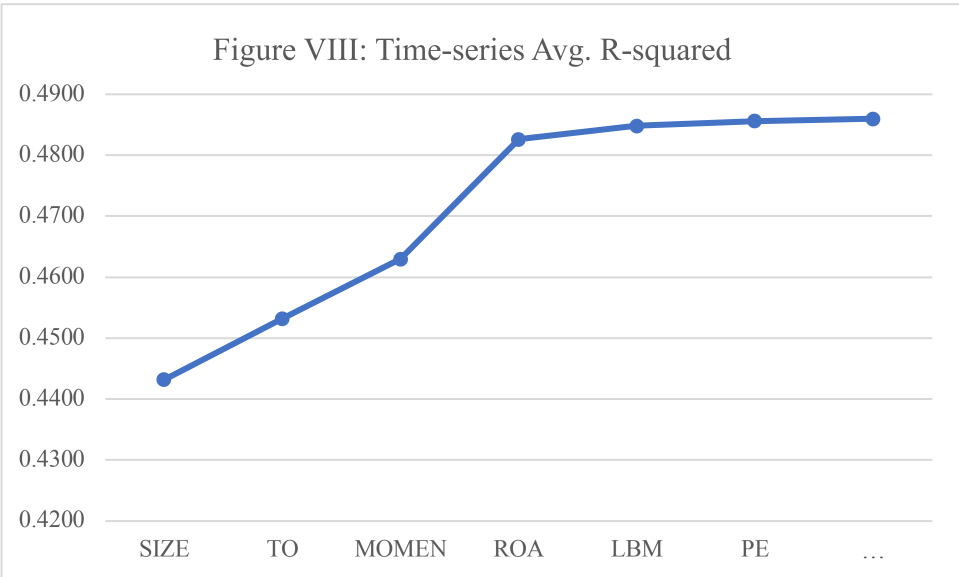
Mingshan Zhou, Jing Lin, and Yunbi An, 2016, Star Analysts, Overreaction, and Synchronicity: Evidence from China and the United States, Financial Management, Volume 46, Issue3, Fall 2017, 797-832.

Appendices

Appendix I

Figure VIII: Firm Characteristic Factors that Determine Analyst Coverage

The figure below contains the time-series average R-squared across regressions of analyst coverage that iteratively added commonly known determinants of analyst coverage, e.g. firm size, share turnover, past performance, ROA, Book to Market ratio and so on. SIZE is the value of market capitalization after logarithm, TO represents share turnover, MOMEN is the cumulative market-adjusted return over the past 12 months at current time point. Our sample consists of 792 firms and 95040 firm-month observations spanning 2008 through 2017.



Appendix II

Figure IX: TCOV Distribution

TCOV Distribution										
	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
0-1	267	205	196	220	258	286	231	192	143	90
1-10	146	160	149	147	132	124	152	169	164	203
10-50	180	190	222	208	186	180	234	271	299	286
50-100	92	237	120	105	103	109	96	103	126	113
>100	107	0	105	112	113	93	79	57	60	100

Figure X: TO Distribution

TO Distribution										
	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
0-300%	43	52	83	218	458	432	365	68	107	349
300%-500%	140	139	150	242	195	228	271	176	200	276
500%-700%	212	174	197	173	72	75	101	224	215	106
700%-900%	225	182	153	79	36	37	32	176	126	41
>900%	172	244	206	79	31	20	23	148	144	20

Figure XI: Size Distribution

Size Distribution										
(million)	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
250-2500	452	169	142	239	219	166	69	2	0	40
2500-4000	114	150	147	164	168	177	142	45	56	107
4000-7000	91	178	182	162	171	160	187	185	191	168
7000-13000	54	135	146	100	99	145	167	218	246	186
>13000	81	160	175	127	135	144	227	342	299	291

Figure XII: Size Distribution

ROA Distribution										
	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
-500%-0%	80	111	68	75	103	84	91	103	91	58
0%-2.5%	173	205	188	178	203	236	231	264	254	230
2.5%-5%	227	231	238	223	215	219	234	203	216	238
5%-8%	140	123	153	156	162	139	136	131	136	164
>8%	172	122	145	160	109	114	100	91	95	102

Figure XIII: Return Distribution

Return Distribution										
	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
-1000%- -20%	526	0	127	530	63	107	3	21	175	310
-20%-0%	60	1	185	150	228	163	35	56	276	215
0%-20%	64	9	180	77	268	177	117	84	224	130
20%-40%	46	24	148	22	132	164	236	126	79	74
>40%	96	758	152	13	101	181	401	505	38	63

Figure XIV: Momen Distribution

Momen Distribution										
	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
-100%- -20%	3	97	41	176	124	80	317	77	92	434
-20%-0%	289	66	186	370	291	176	172	93	298	182
0%-20%	385	86	213	157	224	194	129	143	275	76
20%-40%	73	125	149	69	84	151	86	152	85	43
>40%	42	418	203	20	69	191	88	327	42	57

Appendix III

Table IX: Average Coefficients

The table below shows the residual analyst coverage model estimates using the subsample without data in Dec 2008. SIZE is the value of market capitalization after logarithm, TO represents share turnover, MOMEN is the cumulative market-adjusted return over the past 12 months at current time point. Our sample consists of 792 firms and 94248 firm-month observations.

Table IX: Average Coefficients		
	Mean	t-statistic
INT	-4.495	-62.22
SIZE	0.500	62.92
TO	-0.013	-25.75
MOMEN	0.061	6.76
ROA	1.49	14.13
R^2	0.4820	

Table X : Equally Weighted Average Returns across Portfolios sorted by Total Analyst Coverage

The table below presents equally weighted average one-month-ahead raw returns across 3 portfolios sorted by total analyst coverage, along with a hedge portfolio “High-Low” forming by longing the “High” portfolio at the same time shorting the “Low” portfolio. The subsample use here consisting of 792 firms and 93456 observations.

Table X: Equally Weighted Average Returns across Portfolios sorted by Total Analyst Coverage				
Portfolios	1 (Low)	2	3 (High)	High-Low
EW Average Return	0.0131	0.0115	0.0081	-0.0050
T-statistics	1.4130	1.2663	1.0195	-1.5533

Appendix IV

Figure XV: Average Monthly "High-Low" Strategy Returns Using the Complete Sample

The figure below shows the Avg. one-month holding returns from the strategy that longs the highest quintile and shorts the lowest quintile, where the strategy implements at the end of month m, then hold until the end of month m+1. Our sample consists of 792 firms and 95040 firm-month observations throughout 10 years from 2008 to 2017.