A STUDY OF IDIOSYNCRATIC VOLATILITY IN MUTUAL FUND PERFORMANCE

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERSITY OF HAWAIʻI AT MĀNOA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

INTERNATIONAL MANAGEMENT

MAY 2014

By

Trang Phan

Dissertation Committee:
Wei ("Victor") Huang, Chair
David Hunter, Co-Chair
Qianqiu Liu
Boochun Jung
Timothy Halliday

Keywords: Mutual Fund, Idiosyncratic Volatility, Active Management

ABSTRACT

This research examines idiosyncratic volatility in mutual fund performance. It has two chapters.

In the first chapter, we explore how fund managers have historically invested in stocks with different idiosyncratic volatilities and whether the idiosyncratic volatility of assets in a mutual fund portfolio is correlated with the funds performance. We find that mutual funds underweight stocks with the highest and lowest idiosyncratic volatilities; on average, loadings on high idiosyncratic volatility stocks tend to lower fund performance; and there is no evidence that actively managed funds exploit and profit from the negative relationship between idiosyncratic volatility and stock returns reported in prior studies.

In the second chapter, we present "Residual Correlation" as a unifying measure of mutual fund active management. By decomposing portfolio idiosyncratic volatility into variance and covariance terms, we construct our "Residual Correlation" measure. We propose that skillful fund managers will anticipate positive unsystematic return events. By exposing their portfolios to those events, their portfolios will show correlated returns that are independent of common risk factors. Therefore, active management should be revealed through these correlated asset returns. We show that our measure can identify active management as precisely as various other measures of active management and that it also does so among groups of funds where other measures cannot.

Acknowledgments

I owe my gratitude to many people who have made this dissertation possible and thanks to whom my graduate experience has been one that I will cherish forever.

First and foremost I'd like to thank my dissertation committee chair, Dr. Wei ("Victor") Huang for giving me an opportunity to work on challenging and interesting projects over the past four years. He has always made himself available for help and advice about the dissertation, as well as about how to become a good scholar, a good teacher, and a good colleague in general. I cannot thank him enough for the countless hours he has spent to help with my dissertation, for his invaluable guidance and insightful comments. It has been my great pleasure to work with and learn from such an extraordinary professor.

I would also like to express my deepest gratitude to my co-chair, Dr. David Hunter. Without his extraordinary theoretical ideas, deep knowledge and critical thinking skills, I would never be able to make my dissertation so well developed. Dr. Hunter has taught me so much about forming and developing ideas into good academic papers, as well as about using various databases and research tools.

I am extremely thankful to Dr. Qianqiu Liu for his great care, supports and encouragement, not only with my research but also with other issues in life. I am deeply grateful to Dr. Boochun Jung and Dr. Timothy Halliday for agreeing to serve on my committee, for sparing their invaluable time reviewing the manuscript, and for valuable feedback.

I greatly appreciate the supports, encouragement, and valuable advices from

Dr. Ghon Rhee, Dr. Rosita Chang and Dr. Tung Bui. I am indebted to Dr. Jing Ai, Dr. Jooho Kim, Mr. Benjamin Bystrom, Ms. Judith Mills, Mr. Chris Rachel, Dr. Steven Dawson, Dr. Eric Mais, Dr. James Marsh, Dr. Jack Suyderhoud, Dr. Qimei Chen, Dr. Erica Okada, Dr. Sang-hyop Lee, Dr. Nori Tarui, Dr. Kevin Sun for their valuable help.

Special thanks to my best friend in Hawaii, Hong Vo who has been so encouraging and supporting to me, who has consistently motivated me to do good research and provided me with many valuable research tools and sample computer programs. Hong Vo always keeps abreast of cutting edge literature and discusses many relevant papers to help me extend and strengthen my research papers.

I am also very grateful to my friends Eric Wen, Helen Saar, Taeko Yasutake, Haruyoshi Ito, Dzung Nguyen, Tam Le, Lijuan Zhao, Doug Chun, Nolan Kido, Daegyu Yang, Nguyen Anh Thuc, Chi Diep, Hoang Do and many other friends for their friendship and supports, which made my time in Hawaii much more enjoyable.

I owe my deepest thanks to my family - my husband, my children, my parents and my sister who have always been beside me, strongly supporting me with their unconditional love and encouragement, sharing and going through many hard times with me over these years. A special gratitude goes to my mother in law, who always loved and supported me. Even now when she has passed away, I still feel her continuous love and strong supports.

It is impossible to remember all, and I apologize to those I've inadvertently left out. Thank you all so much.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AS	Active Share
CAPM	Capital Asset Pricing Model
CRSP	Center for Research in Security Prices
CS	Characteristics Selectivity
CT	Characteristics Timing
CV	Covariance Term
IV	Idiosyncratic Volatility
ICI	Industry Concentration Index
MC	Magnitude of Covariance
RC	Residual Correlation
TE	Tracking Error
TNA	Total Net Asset
TRC	Total Return Correlation
VT	Variance Term
WRDS	Wharton Research Data Services

Chapter 1: Idiosyncratic Volatility of Mutual Fund Holdings and Fund Performance

1.1 Introduction

CAPM suggests that idiosyncratic risk is irrelevant in asset pricing because it can be eliminated through diversification. However, idiosyncratic risk has been shown theoretically and empirically to predict stock returns in recent finance literature. Given that majority of investors cannot fully diversify their portfolios and require compensation for holding idiosyncratic risk, some theoretical models predict a positive relation between idiosyncratic risk and stock returns.[1] In contrast, considering various sources of market frictions, several other models suggest a negative relation between idiosyncratic risk and stock returns.[2] The empirical evidences of

[1]See Litner (1965); Levy (1978); Merton (1987) and Lehmann (1990).

[2]Miller (1977) proposes that when investors disagree about stocks true value (as in high idiosyncratic risk stocks), pessimistic investors cannot incorporate their valuation into the stock price because of short-sale constraints. As a result, stocks are overpriced and would have lower returns when overprice is corrected. Constantinides (1984) argues that investors may require lower expected returns for holding stocks with higher idiosyncratic risk since these stocks have a higher tax-timing option value. Sleifer and Vishny (1997) points out that in the presence of limits to arbitrages, "stocks are not rationally priced, and idiosyncratic risk deters arbitrage. In particular,

the relationship between idiosyncratic risk and stock returns are mixed. For example, Fu (2009) shows the positive relationship at the stock level, using a conditional idiosyncratic volatility.[3] On the other hand, Ang, Hodrick, Xing and Zhang (AHXZ - 2006, 2008) show that stocks with higher realized idiosyncratic volatility have lower future returns, even after controlling for size, value, momentum, liquidity, trading volume, and dispersion of analyst forecasts.[4]

Evidence of predictability of idiosyncratic volatility in cross section of stock returns suggests that actively managed equity fund managers may have considered and been able to exploit the significant idiosyncratic volatility premium. We would like to emphasize that the nature and the causes of the relationship between idiosyncratic volatility and stock return is beyond the scope of this study. The issue here is that, given the observed relationship, whether fund managers exploit it.[5]

There is scarce evidence of how funds load on idiosyncratic volatility, let alone

some stocks with high idiosyncratic variance may be overpriced, and this overpricing is not eliminated by arbitrage because shorting them is risky". These volatile overpriced stocks earn a lower expected return. Johnson (2004) proposes a model which values equity as an option and argues that firms with higher idiosyncratic risk may have higher current equity value but lower expected equity given the mixed terminal value of the firm.

[3]Tinic and West (1986) and Makiel and Xu (2002) have documented that portfolios with higher idiosyncratic risk have higher returns.

[4]This result is also confirmed by Zhang (2006), Bali and Cakici (2008), Jiang, Tao and Yao (2009), Huang, Liu, Rhee and Zhang (2010) amongst others.

[5]Specifically, regardless of whether the strong negative relationship between realized idiosyncratic volatility and subsequent stock return demonstrated in AHXZ (2006) is due to return reversal (Huang, Liu, Rhee, and Zhang, 2010) or gambling (Bali, Cakici, Whitelaw, 2010) or any other reason, can mutual funds exploit this negative premium to make profit.

evidence of whether idiosyncratic volatility loading predicts fund performance. The goal of this study is to investigate this issue. Specifically, we look at two different aspects: (i) we explore mutual fund portfolios to understand how fund managers have invested in stocks with different idiosyncratic volatility; (ii) we examine whether funds holding of idiosyncratic volatility is correlated to funds performance in the same way that it is in individual stocks. In other words, we identify how idiosyncratic volatility can be exploited in mutual fund portfolios.

While investigating mutual fund portfolio holdings in the years 1991 and 1992 to find fund preferences for stock characteristics, Falkenstein (1996) infers that mutual fund ownership is concave in idiosyncratic variance, or more accurately mutual funds are averse to very low idiosyncratic volatility. He also claimed that in aggregate, funds have a preference towards securities with higher idiosyncratic volatility. Arguing that residual variance has a high correlation with total variance and adds little information, Falkenstein uses return variance (total variance of monthly returns for the prior 2-5 years) in his tests to infer the results for idiosyncratic variance. Not much has been done since this study. Moreover, what remains unknown is whether fund loading of idiosyncratic volatility can predict fund performance. We will fill this gap by using a direct and commonly accepted measure of idiosyncratic volatility - the realized idiosyncratic volatility (hereinafter referred to as IV).

Following Campbell, Lettau, Makiel and Xu (2001); AHXZ (2006, 2008); and Bali and Cakici (2008), we measure monthly realized idiosyncratic volatility of a stock as the standard deviation of daily residual returns. Our main findings are as follows: Firstly, mutual funds (individually and collectively) underweight the most return-predictable IV quintiles (lowest and highest IV quintiles). On average, 70.84% of mutual funds underweight stocks in low IV quintile while 62.28% of mutual funds underweight stocks in high IV quintile. Moreover, on average 34.1% of funds totally avoid highest IV quintile.

Regardless of the documented strong negative relationship between realized idiosyncratic volatility and stock returns, a majority of mutual funds (73% on average) dont trade on idiosyncratic volatility as revealed by AHXZ. Specifically, only 27.19% of mutual funds trade on idiosyncratic volatility on the direction suggested by AHXZs results, i.e. overweight low idiosyncratic volatility stocks and underweight high idiosyncratic volatility stocks (we call these funds IV anomaly investors, i.e. investors that exploit IV anomaly). Surprisingly, a considerably higher proportion (35.75%) of mutual funds trade on opposite direction, i.e. overweight high idiosyncratic volatility stocks and underweight low idiosyncratic volatility stocks (we call these funds IV anomaly contrarians).

Secondly, in spite of funds avoidance of trading on IV according to AHXZs results, IV anomaly investors tend to have better performance than IV anomaly contrarians, on a risk adjusted basis. It is observable that, on average, funds that load more on high IV stocks tend to underperform. However, there is no evidence that actively managed funds exploit and profit from the negative idiosyncratic volatility premium. Holding higher idiosyncratic volatility stocks tend to hurt funds that are not actively managed.

We proceed as follows. Session 1.2 explains empirical methodology; Section 1.3 describes sample data. Session 1.4 provides results. Session 1.5 concludes.

1.2 Empirical methodology

1.2.1 Realized Idiosyncratic Volatility (IV) of Stocks

Many studies of idiosyncratic risk use expected volatility which is unobservable and must be estimated by sophisticated parametric models. The choice of a parametric model may be essential for volatility forecasting, yet less important for researches investigating historical data to reveal some hidden information. We believe a simple measure of realized idiosyncratic risk would be sufficient, given our purpose in this study is to investigate fund portfolio idiosyncratic volatility to see if it is related to portfolio return and thus can reveal fund manager skill.[6] We simply follow the approach in AHXZ (2006) and Bali and Cakici (2008) to calculate realized idiosyncratic volatility. Specifically, for each month, we run the following Carhart (1997) four factor regression for firms that have more than fifteen daily return observation in that month:[7]

$$r_{t,d}^{i} = \alpha_{t}^{i} + \beta_{t,MTK}^{i} MKT_{t,d} + \beta_{t,SMB}^{i} SMB_{t,d} + \beta_{t,HML}^{i} HML_{t,d} + \beta_{t,UMD}^{i} UMD_{t,d} + \epsilon_{t,d}^{i} \tag{1.1}$$

[6]Campbell, Lettau, Makiel and Xu (200) also use daily data within each month to construct sample variances for that month for a similar reason: *"the choice of a parametric model may be essential for volatility forecasting, it is less important for describing historical movements in volatility, because all models tend to produce historical fitted volatilities that move closely together."* Or as Warren Buffett simply puts: *"In the business world, the rearview mirror is always clearer than the windshield."*

[7]Fama French three factor model was used in AHXZ. We also use this measure in unreported tests, however the result is not different.

where, for day d in month t, $r^i_{t,d}$ is stock is excess return, $MKT_{t,d}$ is the market excess return, $SMB_{t,d}$, $HML_{t,d}$, $UMD_{t,d}$ are the returns on portfolios formed to capture size, book-to-market and momentum effects, respectively, and $\epsilon^i_{t,d}$ is the resulting residual. We use the standard deviation (σ) of daily residuals in month t to measure the individual stocks realized idiosyncratic volatility (IV) for this month.[8]

1.2.2 Classifying funds as IV anomaly investors and IV anomaly contrarians - Matched percentage

In this section, we describe our approach in investigating how funds historically load on IV, and whether or not IV loading predicts fund performance. Given the strong negative IV premium documented in AHXZ (2006, 2008), funds load more heavily on low IV may perform better. The negative IV premium in AHXZ (2006) is driven by the significantly higher return of the stocks in the lowest IV quintile compared to the stocks in the highest IV quintile. Therefore, in order to investigate the predictability of IV loading on fund performance, we divided funds into two groups: *"IV anomaly investors"* which includes funds that overweight stocks in the lowest IV quintile and underweight stocks in the highest IV quintile; *"IV anomaly*

[8]Note that the realized monthly idiosyncratic volatility for a stock is not constructed from the standard deviation of the monthly residuals in Equation (1.1). To measure the monthly idiosyncratic volatility of stock i, we follow French, Schwert, and Stambaugh (1987) and multiply the standard deviation of daily residuals in month t (σ^i_t) by $\sqrt{n^i_t}$ where n^i_t is the number of trading days during month t. Therefore, $IV^i_t = \sqrt{n^i_t}\sigma^i_t$ is stock is realized idiosyncratic volatility in month t.

contrarians" which includes funds that underweight stocks in the lowest IV quintile and overweight stocks in the highest IV quintile.

In order to compare performance of the group of IV anomaly investors and the group of IV anomaly contrarians we can simply take equally weighted return of funds in each group. However, we can also take a different weighting scheme: in the group of IV anomaly investors, we give higher weights to the funds that go further in exploiting AHXZ result; and in the group of IV anomaly contrarians, we give higher weights to the funds that go further in opposite direction of AHXZ result. We define *Matched percentage* for this purpose.

Matched percentage is defined as the minimum of overweight in the lowest IV quintile and underweight in the highest IV quintile for IV anomaly investors; and the minimum of underweight in the lowest IV quintile and overweight in the lowest IV quintile for IV anomaly contrarians. The matched percentage reveals the level that an IV anomaly investor (IV anomaly contrarian) may gain (lose) from deviation of loading on the lowest and highest IV quintile relative to the market benchmark. We predict that higher matched percentage would increase performance for IV anomaly investors and decrease performance for IV anomaly contrarians. Therefore, to compare performance of the group of IV anomaly investors and the group of IV anomaly contrarians, we take matched percentage weighted returns of funds in each group.

1.3 Sample Data

Mutual Fund data come from two sources: (i) The CRSP Survivorship Bias Free Mutual Fund Database includes information on funds characteristics such as fund returns, total net assets, investment objective, turnover ratio, expense ratios and other types of fees. CRSP Mutual Fund Database records different share classes of the same fund as distinct funds; (ii) CDA/Spectrum S12 mutual fund holding database. The CDA database is collected from mutual funds reports filed with SEC and from funds voluntary reports. The two databases are merged using MFLINKS file of Wharton Research Data Services (WRDS).

MFLINKS allows us to match different share classes of the same fund that are recorded as distinct funds in CRSP, into the holdings of the corresponding fund in CDA/Spectrum. We aggregate different share classes of the same fund in CRSP as follows. For age and qualitative characteristics (name, investment objective), we keep the value of the oldest share class. For TNA of the fund, we add up TNAs of its share classes. For other quantitative characteristics, we take the weighted average of other quantitative characteristics (return, expense ratio, management fee, turnover ratio), using the most recent TNA of each share class as weights.[9]

Similar to Kacperczyk, Marcin, Sialm, and Zheng (2008), Kacpercyk et al (2005) and other papers dealing with holding data of mutual funds, we assume that funds carry the same holding from a report date until the next report date (or for 6 months, whichever sooner). We focus on actively managed domestic equity mutual

[9] Also see Kacperczyk, Sialm, Zheng (2006).

funds, eliminating balanced, bond, money market, and international funds.[10] We also exclude funds with TNA less than $5 million because inclusion of smaller funds may cause a survivorship bias problem. Since a stock holding is included only if it can be matched with the CRSP stock file and its Idiosyncratic Volatility can be calculated, we follow Cremers and Petajisto (2009) to include only funds that have more than 67% of value of stock holdings over the fund total net asset. We also require that funds hold ten stocks or more and that fund characteristics (TNA; Age; Turnover; Expenses) are non-missing. After all exclusions, our final sample includes 3921 actively managed equity funds. Our sample period is from 1980 until 2012.

[10] Similary to Huang, Sialm, and Zhang (2011), we select funds with the following Lipper objectives: 'CA', 'CG', 'CS', 'EI', 'FS', 'G', 'GI', 'H', 'ID', 'LCCE', 'LCGE', 'LCVE', 'MC', 'MCCE', 'MCGE', 'MCVE', 'MLCE', 'MLGE', 'MLVE', 'MR', 'NR', 'S', 'SCCE', 'SCGE', 'SCVE', 'SG', 'SP', 'TK', 'TL', 'UT'. When the Lipper code is missing, we select fund with the following Strategic Insights objectives: 'AGG', 'ENV', 'FIN', 'GMC', 'GRI', 'GRO', 'HLT', 'ING', 'NTR', 'SCG', 'SEC', 'TEC', 'UTI', 'GLD', 'RLE'. When both Lipper and SI codes are missing, we select funds with the following Wiesenberger objectives: 'G', 'G-I', 'G-S', 'GCI', 'IEQ', 'ENR', 'FIN', 'GRI', 'HLT', 'LTG', 'MCG','SCG', 'TCH', 'UTL', 'GPM'. If a fund has none of these objective codes but it has a CS policy or has the percentage of common shares in the porfolio between 80%-105%, then the fund will be included. The percentage of common shares is calculated as the time-series average for each fund. We exclude all funds that contain the following words in the fund name: 'INTERNATIONAL', 'GLOBAL', 'BOND', 'BALANCED', 'MONEY MARKET'.

1.4 Empirical Results

In this part, we show empirical results to answer (1) how funds historically load on IV; (2) whether loadings on IV impact fund performance; and (3) whether actively managed funds can exploit the negative IV premium to make profits.

1.4.1 Preferences for IV as revealed by funds holdings

Panel A of table 1.1 shows the summary of the weight that market, as well as all mutual funds put on each IV quintile. At aggregate level, funds as a whole underweight the lowest and highest IV quintile and overweight IV quintiles 2, 3, 4.

Panel B of table 1.1 presents all funds aggregate overweight (underweight) in each IV quintile. On average, mutual funds significantly underload stocks of the lowest (highest) IV quintile by 6.43% (0.27%.) Mutual funds significantly overload stocks of IV quintile 2, 3 and 4 by 2.66%, 2.98% and 1.05% respectively. The minimum percentage underweight in lowest IV is 1.22%. At aggregate level, mutual funds have a strong preference for IV quintile 2, 4 and especially the middle quintile (quintile 3). Only in 5% of months in the sample period, mutual funds overweight IV quintile 3 by less than 1.1%.

In addition to all funds loading on IV at aggregate level, we also look at individual funds' loading on IV. Panel C of table 1.1 shows the summary of percentage of funds that overweight (underweight) stocks in each IV quintile relative to the market weight. On average, 70.84% of funds underweight stocks in the lowest IV quintile and 62.28% of funds underweight stocks in the highest IV quintile. On the other hand, 57.66%; 64.56% and 53.67% of funds overweight stocks in IV quintiles 2; 3 and 4 respectively.

Moreover, the minimum percentage of funds underweight lowest IV quintile is 59.35%. I.e. every month in the sample period, there are at least 59.35% of funds underweighting stocks in the lowest IV quintile. Also, only in 10% of months in the sample period, the percentage of funds underloading the highest IV quintile is less than 50.12%. That means in most of sample period, more than half of funds underloading the highest IV quintile. Furthermore, only in 5% of months in the sample period, the percentage of funds that overload the middle IV quintile (quintile 3) is less than 57.32%. Thus in most of sample period, more than half of funds overload the middle IV quintile.

Panel D of table 1.1 shows the percentage of stocks in each IV quintile collectively held by actively managed equity funds, in terms of number of stocks held and total market value of stocks held. With regards to the market value of stocks in each IV quintile that is held by funds, funds hold 7.8% of market value of stocks in IV quintile 1; 11% of market value of stocks in IV quintile 3 and 7.7% of market value of stocks in IV quintile 5.

Considering the number of stocks in each IV quintile that is held by funds, we see that on average, funds hold 66.5% of stocks in the lowest IV quintile, 57.8% of stocks in the highest IV quintile and 83% of stocks in the 3rd IV quintile. In other word, there are 42.2% of stocks in the highest IV quintile and 33.5% of stocks in the lowest IV quintile that are totally avoided by funds.

Panel E of table 1.1 exhibits the percentage of funds that hold stocks in each

IV quintile. On average, 97.6%; 98.2% and 65.9% of funds hold stocks in IV quintile 1, 3 and 5 respectively. Its worth pointing out that 34.1% of funds totally avoid (did not hold any) stocks in the highest IV quintile, and 10% of funds avoid stocks in IV quintile 4. This might be because uninformative firms would want to stay away highly informative stocks (high IV stocks) which are risky and not rewarded.

Related to the issues analyzed in table 1, Falkenstein (1996) was the first to show that funds are averse to stocks with low idiosyncratic volatility; and mutual fund ownership is concave in idiosyncratic variance. He claims that in aggregate, funds have preferences towards securities with higher volatility. He uses two years of holding data (1991, 1992) and return variance (total variance of monthly returns for the period 2-5 years). Arguing that residual variance has a high correlation with total variance and adds little information, Falkenstein conducts tests with return variance and infer the results for idiosyncratic volatility. There is no evidence of significant relationship between return variance and future returns, so using residual variance would not help to investigate predictability of holding idiosyncratic risk on funds performance.

With results demonstrated in table 1.1, we confirm Falkensteins finding for a much longer sample period (1980 to 2012). More importantly, we use realized idiosyncratic volatility (IV), a measure that would allow us to explore whether funds can exploit loading on IV to improve performance, given the strong negative IV premium documented in AHXZ (2006, 2008.) In addition, we find that funds are also averse to stocks with high idiosyncratic volatility and that funds tend to have a preference for stocks in the middle IV quintiles (especially IV quintile 3).

This trend is seen in both individual funds and aggregate of all funds. This trend is observed in both percentage of stocks in each IV quintile that is held by funds and percentage of funds that hold stocks in each IV quintile.

1.4.2 Impact of IV loading on funds performance

We study whether or not IV loading impacts funds performance. This is done in two different ways. Firstly, we investigate if funds can apply the negative relationship of stocks IV and return as documented in AHXZ (2006) to create abnormal returns. Given the observed negative IV premium in the stock market, a fund that spreads its holdings to a large number of high idiosyncratic risk assets will have a high weighted average of idiosyncratic volatility, which should result in low performance.

In this section, we examine the relative performance of funds in quintiles of weighted average IV of assets in the fund portfolio. Every month, we sort all mutual funds into five quintile portfolios based on weighted average IV. For each quintile portfolio, we compute the equally weighted average next month gross return (reported net return plus expense ratio). The difference of equally weighted average returns of the top and the bottom quintile portfolios is also calculated. We then report the time-series average of these portfolio returns (and of the return difference of top and bottom quintile); the CAPM alpha; Fama French three factor alpha and the four factor alpha.

Panel A of table 1.2 exhibits that the difference in performance of top and bottom quintile portfolios sorted by weighted average IV is not significant. The difference in the alphas between top and bottom portfolio tend to be negative and might be driven by the negative relationship between idiosyncratic risk and subsequent month return in the stock market as demonstrated by AHXZ (2006). Huang, Liu, Rhee and Zhang (2010) show that AHXZ (2006) result is driven by the return reversal observed in the stock market. In Panel B of table 1.2, we report the performance in the fourth month of the quintiles portfolio sorted by the weighted average IV. The result is similar to that in Panel A: the difference in the alphas between top and bottom quintile portfolio tend to be negative but not significant.

Regardless of the result demonstrated by AHXZ (2006) that assets with high idiosyncratic volatility have abysmally low average returns, the insignificant relationship between weighted average IV and fund performance is not surprising. This suggests that funds do not randomly select high idiosyncratic risk assets (which may lead to underperformance), it's likely that they selectively pick up the assets they have superior information.

Secondly, we specifically look into two groups of funds: (1) funds that overload stocks in the lowest IV quintile and underload stocks in the highest IV quintile. (We call these funds IV anomaly investors); (2) funds that underload stocks in the lowest IV quintile and overload stocks in the highest IV quintile. (We call these funds IV anomaly contrarians). We wanted to conduct this test because Bali and Cakici (2008) show that AHXZ (2006) result is driven by very low return of the highest idiosyncratic volatility assets, thus the stocks in the lowest IV quintile significantly outperform stocks in the highest IV quintile.

Panel A of table 1.3 presents the summary statistics of the percentage of IV anomaly investors; IV anomaly contrarians; as well as the percentages of funds that overweight (underweight) stocks in both lowest and highest IV quintile. On average, only 27.19% of mutual funds trade on idiosyncratic volatility on the direction suggested by AHXZs results, i.e. overweight low idiosyncratic volatility stocks and underweight high idiosyncratic volatility stocks. Surprisingly, a considerably higher proportion (35.75%) of mutual funds trade on opposite direction, i.e. overweight high idiosyncratic volatility stocks and underweight low idiosyncratic volatility stocks.

Panel B of table 1.3 reports the holding portfolio returns[11] (both raw and risk adjusted returns) of IV anomaly investors compared to IV anomaly contrarians. IV anomaly investors seem to perform better on a risk adjusted basis with a significant CAPM alpha of 0.13% per month. Carhart (1997) four factor alpha and Fama French (1993) three factor alpha are also positive with t-values considerably higher than those of IV anomaly contrarians.

To further show the better performance of IV anomaly investors relative to IV anomaly contrarians, we calculate weighted average holding returns of each group, with the weight being the matched percentage of overweight (underweight) in IV 1

[11]This test has also been done for mutual funds raw returns, however the result seems somewhat weaker than the result obtained using portfolio holding returns. This suggests that classification of IV anomaly contrarians and IV anomaly investors is independent with funds skills. In other word, funds just happen to be IV anomaly investors or contrarians. Regardless, the predictability of funds loading on IV in funds performance tends to exist.

stocks and underweight (overweight) in IV 5 stocks for IV anomaly investors (IV anomaly contrarians). Specifically, matched percentage = min (overweight in IV 1, underweight in IV 5) for IV anomaly investors, and matched percentage = min (underweight in IV 1, overweight in IV5) for IV anomaly contrarians.

Panel C of table 1.3 exhibits the match weighted returns (and CAPM, Fama French (1993) three factor and Carhart (1997) four factor alphas) of IV anomaly investors relative to IV anomaly contrarians. Again, the results confirm that IV anomaly investors tend to have superior performance as compared to IV anomaly contrarians with significant CAPM alpha (0.15% per month) and Fama French (1993) three factor alpha (0.08% per month).

1.4.3 Impact of IV loading on funds performance after controlling for active management

In the previous section, we showed that funds loading more on low IV tend to perform better (but not significant). Impact of IV loading on funds' performance may depend on whether funds are actively managed. If funds are not actively managed, loading more on high IV may lead to underperformance. However, if funds are actively managed, it may selectively pick up stocks with high IV in order to exploit and profit from firm-specific information.

In Panel A of table 1.4, we first sort funds into three portfolios based on Industry Concentration Index (ICI)[12]. Next, we further sort funds into five quintiles

[12]Industry Concentration Index is defined as in Kacpercyk et al (2005) $ICI_t = \Sigma_{j=1}^{10}(w_{j,t} - \overline{w}_{j,t})^2$ where $w_{j,t}$ is the weight the fund loads on industry j, and $\overline{w}_{j,t}$ is the weight of industry j in the

based on Weighted Average IV and report the four factor alphas (calculated based on equally weighted gross return of funds in each portfolio). We observe that if funds are not actively managed (having low ICI), loading more on high IV stocks may hurt fund performance: in this low ICI group of funds, funds in the highest weighted average IV quintile portfolio significantly underperformed funds in the lowest weighted average IV quintile portfolio. This is predictable, because if a fund hold stocks with more firm specific information, yet unable to exploit the information to make profit, it is likely that the fund would underperform. If funds are actively managed, they may pick up high IV stocks based on their information/skills, so funds holding high IV stocks may not underperform.

Panels B and C of table 1.4 show similar stories, i.e. loading more on high IV stocks may hurt fund performance when the funds are not actively managed (having low AS or having high R^2)[13]. In all, we did not find an evidence that actively managed funds may profit from exploiting the negative IV premium documented in AHXZ (2006, 2008).

---

stock market.

[13]Active Share is provided by Cremers and Petajisto (2009) and is defined as $AS = \frac{1}{2}\Sigma_{i=1}^{N}|w_{fund,i} - w_{index,i}|$ where $w_{fund,i}$ and $w_{index,i}$ are the weights of asset i in the holdings portfolio and in the index, and the sum is taken over the universe of all assets. R^2 is defined as in Amihud and Goyenko (2013), obtained from regressions of fund monthly gross return (reported net return plus expense ratio) over previous 24 months on the Four Factor Model.

1.5 Conclusion

Idiosyncratic risk has been shown theoretically and empirically to predict future return in stocks. AHXZ (2006, 2008) demonstrate that stocks with low realized idiosyncratic volatility (IV) significantly outperform stocks with high IV. This result has been confirmed in other papers. This chapter studies mutual fund investment in IV and investigates whether funds can exploit IV to improve performance.

Our main findings include (1) Funds underweight high IV and low IV stocks; In addition, funds totally avoid a considerable number of high IV stocks. Funds tend to overweight stocks in the middle IV quintiles (especially IV quintile 3); (2) Funds may be able to take advantage of the negative relationship between stock IV and stock return as stated in AHXZ (2006) to have better performance. On average, loadings on higher IV stocks lower fund performance. Furthermore, funds that overweight stocks in the bottom IV quintile and underweight stocks in the top IV quintile tend to have better performance than funds that underweight high IV stocks and overweight low IV stocks. However, we did not find any evidence that actively managed funds exploit and profit from this negative IV premium.

Given the observed relationship between stocks IV and return as documented in AHXZ (2006), it is understandable why funds underweight high IV stocks. However, it is puzzling to us why funds have an aversion to low IV stocks. In our future research, we would like to study possible motivations for funds to invest in high or low IV stocks. We would also like to study the evolvement of funds loadings on high/low IV over time.

Table 1.1: Preferences for IV

Panel A: Statistics of market and aggregate funds' weight of each IV quintile
Every month, total market weight and all mutual funds' weight of each stock IV quintile are calculated. This table reports the time-series statistics of market weight and funds' weight of each IV quintile. The sample period is from 1980 to 2012.

	Market weight				All funds' weight			
IV Quintile	Avg	Std	Min	Max	Avg	Std	Min	Max
1	46.17%	11.77%	17.83%	70.40%	39.77%	11.23%	13.92%	65.78%
2	30.96%	6.22%	19.20%	48.98%	33.63%	5.28%	23.11%	49.08%
3	14.26%	4.19%	6.63%	27.14%	17.22%	4.39%	6.97%	28.35%
4	6.41%	2.41%	2.39%	13.75%	7.46%	3.02%	2.39%	16.67%
5	2.19%	1.19%	0.62%	8.92%	1.92%	1.35%	0.21%	7.51%

Panel B: Statistics of aggregate funds overweight (underweight) of each IV quintile
Every month, the aggregate overweight (underweight) of all funds holdings in each stock IV quintile is calculated. This table reports the time series statistics of the annual aggregate overweight (underweight). The sample period is from 1980 to 2012.

IV Quintile	Mean	Median	Std Dev	Min	Max	5th Pctl	10th Pctl	90th Pctl	95th Pctl	t Value
1	-6.43%	-6.42%	2.06%	-12.94%	-1.22%	-9.90%	-9.25%	-3.83%	-3.12%	-60.02
2	2.66%	2.69%	1.94%	-4.36%	8.68%	-0.58%	0.21%	5.15%	5.83%	26.41
3	2.98%	3.01%	1.13%	-2.26%	7.05%	1.10%	1.59%	4.39%	4.63%	50.56
4	1.05%	0.92%	0.93%	-0.68%	4.53%	-0.25%	0.00%	2.41%	2.89%	21.82
5	-0.27%	-0.31%	0.41%	-1.70%	1.20%	-0.94%	-0.67%	0.22%	0.50%	-12.46

Panel C: Summary of funds' overweight (underweight) percentage in each IV quintile

This table reports the summary of percentage of funds that overload (underload) in each IV quintile. Every month, the percentage of funds that overload (OW=1) or underload (OW=-1) in each IV quintile is calculated. A fund overloads (underloads) in an IV quintile when the holding weight of this IV quintile is more (less) than that of the benchmark market. The table reports the time-series statistics for the sample period from 1980 to 2012.

IV Quintile	OW	AvgProp	StdProp	MinProp	MaxProp	5th Pctl	95th Pctl	10th Pctl	90th Pctl
1	-1	70.84%	4.72%	59.35%	86.36%	63.25%	78.55%	64.89%	76.95%
1	1	29.16%	4.72%	13.64%	40.65%	21.45%	36.75%	23.05%	35.11%
2	-1	42.34%	10.04%	22.49%	70.64%	26.94%	58.01%	28.99%	55.37%
2	1	57.66%	10.04%	29.36%	77.51%	41.99%	73.06%	44.63%	71.01%
3	-1	35.44%	4.77%	22.98%	59.34%	28.01%	42.68%	29.65%	41.70%
3	1	64.56%	4.77%	40.66%	77.02%	57.32%	71.99%	58.30%	70.35%
4	-1	46.33%	6.00%	31.51%	60.53%	36.32%	55.51%	38.14%	53.93%
4	1	53.67%	6.00%	39.47%	68.49%	44.49%	63.68%	46.07%	61.86%
5	-1	62.28%	8.69%	39.32%	83.63%	47.11%	75.69%	50.12%	73.30%
5	1	37.72%	8.69%	16.37%	60.68%	24.31%	52.89%	26.70%	49.88%

Panel D: Percentage of stocks in each IV quintile held by funds (in terms of number of stocks and market value of stocks)

Each month, we calculate the number and market value of stocks in each IV quintile that collectively held by funds. We also calculate the number and market value of stocks in each IV quintile for the benchmark market index. Next, we calculate the percentage of stocks in each IV quintile collectively held by actively managed equity funds in terms of number of stocks held and total market value of stocks held. This table shows the time series statistics for percentage of stocks in each IV quintile collectively held by funds for the sample period from 1980 to 2012.

IV quintile	Variable	Mean	Median	Std Dev	Minimum	Maximum
1 (Lowest)	Percentages of number of stocks held	66.50%	70.80%	12.80%	21.70%	87.90%
	Percentages of market value of stocks held	7.80%	7.60%	4.40%	0.20%	14.20%
2	Percentages of number of stocks held	83.00%	83.00%	7.00%	46.40%	93.00%
	Percentages of market value of stocks held	9.80%	10.30%	5.40%	0.30%	18.30%
3	Percentages of number of stocks held	83.00%	86.70%	9.30%	43.90%	95.90%
	Percentages of market value of stocks held	11.00%	12.20%	6.10%	0.40%	20.30%
4	Percentages of number of stocks held	76.30%	81.50%	14.00%	29.60%	95.10%
	Percentages of market value of stocks held	10.50%	10.70%	6.00%	0.40%	19.80%
5 (Highest)	Percentages of number of stocks held	57.80%	57.60%	21.00%	12.60%	89.50%
	Percentages of market value of stocks held	7.70%	5.60%	5.40%	0.30%	19.10%

Panel E: percentage of funds that hold stocks in each IV quintile

Every month, we calculate the percentage of funds that hold stocks in each IV quintile. This table reports the time series statistics of the percentage. The sample period is from 1980 to 2012.

IV quintile	Mean	Median	Std Dev	Minimum	Maximum
1 (Lowest)	97.60%	97.80%	1.30%	92.70%	100.00%
2	99.50%	99.60%	0.50%	96.50%	100.00%
3	98.20%	98.80%	1.60%	91.30%	100.00%
4	90.00%	92.00%	7.00%	65.90%	99.20%
5 (Highest)	65.90%	66.30%	14.10%	32.30%	93.50%

Table 1.2: Performance (before expenses) of quintile portfolios sorted by weighted average IV

This table presents equally weighted monthly gross returns (reported net return plus expense ratio) and risk adjusted returns (CAPM alpha; Fama French (1993) three factor alpha; Carhart (1997) four factor alpha) for quintile portfolio of funds, sorted by Weighted Average IV of assets in the fund portfolio. Portfolios are rebalanced monthly, basing on the sorting variables calculated in the previous month (Panel A) or fourth month before (Panel B). The table also reports the difference in the returns and risk adjusted returns between the top and bottom quintiles. Robust Newey-West (1987) t-statistics are reported in parentheses. *** 1% significance, ** 5% significance, * 10% significance. Sample period is January 1980 - November 2010.

Panel A: Performance in the next month of quintile portfolio of funds

	Quintile portfolio sorted by weighted average IV					
Variable	1	2	3	4	5	5-1
Average Gross Ret	1.04***	0.99***	1.02***	1.09***	1.04***	0.01
	(4.85)	(4.08)	(3.88)	(3.78)	(2.88)	(0.03)
CAPM Alpha	0.16**	0.03	0.02	0.05	-0.12	-0.28
	(2.24)	(0.71)	(0.47)	(0.71)	(-0.87)	(-1.51)
Fama French Three Factor Alpha	0.08	-0.01	0	0.04	-0.03	-0.11
	(1.54)	(-0.16)	(0.05)	(0.92)	(-0.38)	(-1.07)
Four Factor Alpha	0.09*	0	0	0.04	-0.02	-0.11
	(1.74)	(0.04)	(0.09)	(0.74)	(-0.26)	(-1.08)

Panel B: Performance in the fourth month of quintile portfolio of funds

	Quintile portfolio sorted by weighted average IV					
Variable	1	2	3	4	5	5-1
Average Gross Ret	1.02***	0.98***	1.04***	1.06***	1.05***	0.03
	(4.83)	(4.15)	(4.02)	(3.72)	(2.95)	(0.13)
CAPM Alpha	0.15**	0.04	0.05	0.03	-0.1	-0.25
	(2.13)	(0.93)	(1.13)	(0.42)	(-0.75)	(-1.40)
Fama French Three Factor Alpha	0.07	0.01	0.03	0	-0.05	-0.12
	(1.46)	(0.15)	(0.62)	(0.10)	(-0.62)	(-1.21)
Four Factor Alpha	0.07	0.02	0.03	0	-0.05	-0.11
	(1.34)	(0.48)	(0.53)	(-0.03)	(-0.59)	(-1.16)

Table 1.3: IV anomaly investors vs. IV anomaly contrarians

Panel A: Percentage of IV anomaly investors, contrarians and other funds
Each month, we calculate the percentage of funds that overweight stocks in the lowest IV quintile and underweight stocks in the highest IV quintile (IV anomaly investors); the percentage of funds that underweight stocks in the lowest IV quintile and overweight stocks in the highest IV quintile (IV anomaly contrarians) and funds that overweight (underweight) stocks in both the highest and the lowest IV quintile. This table reports the time-series statistics for the sample period of 1980 to 2012.

Group of funds	Average	Std.	Min	Max
IV anomaly investors (overweight IV1 and underweight IV5)	27.19%	4.95%	12.88%	39.29%
IV anomaly contrarians (overweight IV5 and underweight IV1)	35.75%	8.19%	14.16%	57.85%
Funds overweight both IV1 and IV5	1.98%	1.29%	0.11%	8.90%
Funds underweight both IV1 and IV5	35.09%	7.10%	18.79%	53.66%

Panel B: Equally weighted performance of IV anomaly investors relative to IV anomaly contrarians
Each month, we calculate equally weighted holding return of portfolio of IV anomaly investors (funds that overweight stocks in the lowest IV quintile and underweight stocks in the highest IV quintile); of IV anomaly contrarians (underweight stocks in the lowest IV quintile and overweight stocks in the highest IV quintile). The difference in the monthly return of the two portfolios is also calculated. The second row shows the time-series average holding return; The fourth, sixth and eighth rows report CAPM alpha; Carhart (1997) alpha and Fama French (1993) alpha. Robust Newey-West (1987) t-statistics are reported in parentheses. The sample period is from 1980 to 2012. *** 1% significance; ** 5% significance; * 10% significance.

Type of return	IV anomaly contrarians	IV anomaly investors	Difference
Average ret	1.08***	1.05***	-0.03
	(3.12)	(4.56)	(-0.16)
CAPM alpha	-0.06	0.13**	0.19
	(-0.52)	(2.1)	(1.22)
Fama French alpha	0	0.07	0.07
	0	(1.54)	(0.89)
Carhart alpha	-0.01	0.06	0.07
	(-0.11)	(1.31)	(0.87)

Panel C: Matched percentage weighted performance of IV anomaly investors relative to IV anomaly contrarians

Each month, we calculate matched percentage weighted holding return of portfolio of IV anomaly investors (funds that overweight stocks in the lowest IV quintile and underweight stocks in the highest IV quintile); of IV anomaly contrarians (funds that underweight stocks in the lowest IV quintile and overweight stocks in the highest IV quintile). Matched percentage is defined as the minimum of overweight in IV1 and underweight in IV5 for IV anomaly investors and the minimum of underweight in IV1 and overweight in IV5 for IV anomaly contrarians. The difference in the monthly holding return of the two portfolios is also calculated. The second row shows the time-series average holding return; The fourth, sixth and eighth rows report CAPM alpha; Carhart (1997) alpha and Fama French (1993) alpha. Robust Newey-West (1987) t-statistics are reported in parentheses. The sample period is from 1980 to 2012. *** 1% significance; ** 5% significance; * 10% significance.

Type of return	IV anomaly contrarians	IV anomaly investors	Difference
Average ret	1.05***	1.06***	0.01
	(2.73)	(4.66)	(0.04)
CAPM alpha	-0.14	0.15**	0.29
	(-0.88)	(2.2)	(1.37)
Fama French alpha	-0.05	0.08*	0.13
	(-0.58)	(1.68)	(1.19)
Carhart alpha	-0.01	0.07	0.09
	(-0.14)	(1.44)	(0.75)

Table 1.4: Average performance of portfolio sorted on weighted average IV, controlling for active management measures (Industry Concentration Index, Active Share and R^2)

This table reports four factor alphas (based on equally weighted monthly raw return for different portfolios of mutual funds). Each month, we first sort funds into three portfolios on the basis of lagged active management measures (Industry Concentration Index in Panel A; Active Share in Panel B; and R^2 in Panel C). Then, within each fund portfolios, funds are sorted into five portfolios based on weighted average IV. Therefore, they represent risk adjusted performance of weighted average IV quintile portfolios controlling for other active management measures. Industry Concentration Index is defined as in Kacpercyk et al (2005) $ICI_t = \Sigma_{j=1}^{10}(w_{j,t} - \overline{w}_{j,t})^2$ where $w_{j,t}$ is the weight the fund loads on industry j, and $\overline{w}_{j,t}$ is the weight of industry j in the stock market. Active Share is provided by Cremers and Petajisto (2009) and is defined as $AS = \frac{1}{2}\Sigma_{i=1}^{N}|w_{fund,i} - w_{index,i}|$ where $w_{fund,i}$ and $w_{index,i}$ are the weights of asset i in the holdings portfolio and in the index, and the sum is taken over the universe of all assets. R^2 is defined as in Amihud and Goyenko (2013), obtained from regressions of fund monthly gross return (reported net return plus expense ratio) over previous 24 months on the Four Factor Model. Robust Newey-West (1987) t-statistics are reported in parentheses. *** 1% significance, ** 5% significance, * 10% significance. Sample period is January 1980 - November 2012.

Panel A: Four factor alpha of quintile portfolio sorted by weighted average IV, controlling for Industry Concentration Index (ICI)

	Quintile portfolio sorted by weighted average IV					
ICI Portfolio	1 Low	2	3	4	5 High	5-1
Low ICI	0.02	0	-0.03	0.01	-0.08	-0.10*
	(0.70)	(0.06)	(-0.91)	(0.17)	(-1.42)	(-1.77)
Medium ICI	0.02	0.02	-0.02	0	-0.1	-0.12
	(0.39)	(0.38)	(-0.40)	(-0.03)	(-1.37)	(-1.40)
High ICI	0.16*	0.09	0.14*	0.08	0.06	-0.1
	(1.82)	(1.26)	(1.81)	(0.84)	(0.43)	(-0.63)

Panel B: Four factor alpha of quintile portfolio sorted by weighted average IV, controlling for Active Share

	Quintile portfolio sorted by weighted average IV					
AS Portfolio	1 Low	2	3	4	5 High	5-1
Low AS	0.05	0.05*	0.01	0	-0.01	-0.05
	(1.31)	(1.94)	(0.44)	(0.07)	(-0.10)	(-0.69)
Medium AS	0.09	-0.01	-0.02	0	0.05	-0.04
	(1.49)	(-0.19)	(-0.32)	(0.04)	(0.51)	(-0.30)
High AS	0.1	0.1	0.07	0.01	0.07	-0.03
	(1.29)	(1.28)	(0.91)	(0.08)	(0.61)	(-0.24)

Panel C: Four factor alpha of quintile portfolio sorted by weighted average IV, controlling for R^2

	Quintile portfolio sorted by weighted average IV					
R^2 Portfolio	1 Low	2	3	4	5 High	5-1
Low R^2	0.11	0.09	0.02	0.07	0.03	-0.08
	(1.53)	(1.06)	(0.27)	(0.71)	(0.22)	(-0.64)
Medium R^2	0.05	-0.01	0.01	0.02	-0.06	-0.1
	(0.95)	(-0.30)	(0.17)	(0.38)	(-0.64)	(-0.95)
High R^2	0.05	0.01	-0.04	-0.04	-0.06	-0.11
	(1.19)	(0.51)	(-1.40)	(-0.95)	(-0.68)	(-1.02)

Chapter 2: Residual Correlation and Mutual Fund Performance

2.1 Introduction

In this chapter, we propose an intuitive and parsimonious measure of active management that we call “Residual Correlation”. “Residual Correlation” not only confirms and extends the findings of other existing measures but also can identify active management and predict fund performance in groups of funds where active management previously has not been investigated. We claim that fund managers make investments that they expect will produce superior returns given their anticipation of positive idiosyncratic return events. These investments will generate correlated asset returns that are independent of common risk factors. We study these correlated asset returns and show that they identify active management.

There are two dimensions of active management, stock selection which involves picking individual stocks that are expected to outperform their peers; and factor timing which involves time-varying bets on systematic risk factors. Many recent measures of active management, such as the “Industry Concentration Index”[1] of

[1] Kacpercyk et al (2005) define ICI as follow: Each stock is assigned to one of the 10 industries (as classified in Kenneth French Web page). Every month, ICI is calculated for each fund: $ICI_t = \Sigma_{j=1}^{10}(w_{j,t} - \overline{w}_{j,t})^2$ where $w_{j,t}$ is the weight the fund loads on industry j, and $\overline{w}_{j,t}$ is the weight of

Kacperczyk et al (2005), "Active Share"[2] and "Tracking Error"[3] of Cremers and Petajisto (2009), and "R^2"[4] of Amihud and Goyenko (2013) are similar in that their measures successfully define stock selection of actively managed mutual funds and show that stock selection can earn abnormal returns. Stock selection can be measured with idiosyncratic volatility from a multifactor benchmark model. We extend the existing literature by showing that the idiosyncratic volatility of a fund portfolio can be decomposed into three components that reflect different aspects of active management. We separately investigate the impact of each component on fund performance.

"Residual Correlation" of mutual funds has increased sharply over our sample period (1980 to 2012). In particular, this increase is found only in the highest quintile of funds classified by this measure. "Residual Correlation" is also highly skewed, with the mass of funds having very low "Residual Correlation" (i.e. most funds are not actively managed). This is consistent with skewness of other active management measures, such as "Industry Concentration Index" and R^2. "Residual

---

industry j in the stock market.

[2] "Active Share" is defined as the percentage of a fund's portfolio holdings that differ from the fund's benchmark index. Specifically, $AS = \frac{1}{2}\Sigma_{i=1}^{N}|w_{fund,i} - w_{index,i}|$ where $w_{fund,i}$ and $w_{index,i}$ are the weights of asset i in the fund portfolio and in the index, and the sum is taken over the universe of all assets. See Cremers and Petajisto (2009) for the intuitive economic interpretation of "Active Share".

[3] Tracking error is defined as the time-series standard deviation of the difference between a fund return $R_{fund,t}$ and its benchmark index return $R_{index,t}$: Tracking error $= \sigma_{(R_{fund,t} - R_{index,t})}$.

[4] Stock Selection (or Selectivity) is measured by $1 - R^2 = \frac{RMSE^2}{VARIANCE} = \frac{RMSE^2}{SystematicRisk^2 + RMSE^2}$. $RMSE^2$ is the idiosyncratic volatility.

Correlation" is distinct from existing measures of active management, because it directly identifies precise components within idiosyncratic volatility and links them to active management returns. By decomposing idiosyncratic volatility into three components, we can identify how much each component of idiosyncratic volatility relates to current measures of active management.

First, we decompose a fund portfolio idiosyncratic volatility into two components, the "Variance Term" which contains variances of constituent asset residual returns, and the "Covariance Term" which contains covariances of asset residual returns. Next, we split the "Covariance Term" into the product of the "Magnitude of Covariances" which contains the sum of weighted pairwise asset covariances assuming perfect correlations, and the "Residual Correlation" which is the weighted average of pairwise correlations of asset residual returns. With this decomposition, the portfolio idiosynratic volatility[5] ($\sigma^2_{\epsilon_p}$) can be written as the sum of "Zero Correlation Idiosyncratic Volatility" ($\sigma^2_{\epsilon_p,\rho=0}$), which is the idiosyncratic volatility of the same portfolio assuming no residual correlations, and "Perfect Correlation Idiosyncratic Volatility" ($\sigma^2_{\epsilon_p,\rho=1}$), which is idiosyncratic volatility of the same portfolio assuming perfect residual correlations. These two terms are scaled by "(1-Residual Correlation)" and "Residual Correlation" respectively. Therefore, "Residual Correlation" is a measure representing where portfolio idiosyncratic volatility lies in the line connecting two extremes, $\sigma^2_{\epsilon_p,\rho=0}$ and $\sigma^2_{\epsilon_p,\rho=1}$. $\sigma^2_{\epsilon_p}$ would be closer to $\sigma^2_{\epsilon_p,\rho=1}$ ($\sigma^2_{\epsilon_p,\rho=0}$) if "Residual Correlation" is large (small).

[5]Idiosyncratic volatility is defined as the variance of residual returns from the regression of fund returns on a multi factor benchmark model.

"Residual Correlation" captures the return effects of fund managers' bets on correlated idiosyncratic returns of individual assets. As suggested by Cremers and Petajisto (2009), broad systematic factors may either be too efficiently priced or too difficult for fund managers to predict, so funds that focus on factor bets generally do not outperform. However, fund managers may still be skillful in predicting and/or reacting quickly to idiosyncratic events, which impact only select groups of assets. By anticipating future events, active fund managers will increase their exposure to the stocks that they expect will be favorably impacted by those events. Those stocks will show correlated residual returns as they simultaneously earn abnormally positive returns.

As one example, on September 22, 2010, Abbott Labs announced a Similac Bottle contamination recall. As a consequence of this recall announcement, Abbott Labs' sales dropped sharply while sales of the firm's competitors such as Mead Johnson and Nestle increased because of worldwide customers' switch. The stock market seemed to have a slow reaction upon this announcement: the returns on the month following the recall announcement were 2.94% for Abbott Labs; 4.94% for Johnson Mead; and 0.48% for Nestle. One month later, when Abbott Labs announced its earning, the stock market reacted strongly to Abbott Labs' negative earning surprise: the returns on the following month were -7.2% for Abbott Labs; 12.98% for Mead Johnson; and 5.09% for Nestle. Mead Johnson and Nestle had correlated returns because they acted alike in response to Abbott Labs' firm-specific events. These correlated returns were independent of systematic risk factors or industry designation (Abbott Labs; Mead Johnson and Nestle are in Pharmaceutical; Pediatric Nutrition; and Food Processing industry respectively). Skillful funds managers may have reacted more quickly than the general stock market and sold short Abbott Labs while purchasing Mead Johnson and Nestle and make a superior profit.[6] The high residual return correlation of Mead Johnson and Nestle increased those funds' "Residual Correlation". Similar examples could be Walgreen strikes due to union's problems; Steve Job's health problem; Giorgio Armani's toxic perfumes. These events, while having a negative impact on a single firm, may have a positive impact on that firm's competitors.

If mutual funds abnormal returns result from correlated idiosyncratic returns of individual assets, we would expect a positive relationship between mutual funds' "Residual Correlation" and future performance. If these correlated idiosyncratic returns are due to skill, we should expect to find persistent correlated idiosyncratic returns in select groups of skilled mutual funds. If the correlated bets are due to fund's uninformed speculation, we would not expect to find a persistent relationship between "Residual Correlation" and future fund performance.

We not only investigate the performance predictability of "Residual Correlation" among broad ranges of funds, but we also examine this performance predictability among specific groups of funds. Active management has not been investigated among these funds, because existing measures, by construction, have been

[6] We searched the CRSP Mutual Funds database for examples of funds that might have profited from this particular event, and identified the AXA Premier VIP Trust - Multimanager Large Cap Value Portfolio. This fund is actively managed, and held a much larger portfolio allocation to Mead Johnson than they did to Abbott Labs from October 2010 and thereafter.

unable to do so. These groups include: index funds; closet indexes; and sector funds. Since “Residual Correlation” does not depend on identification of arbitrary industry or benchmark measures, it is particularly well suited to identify active management among these more specialized groups of funds.

Our main findings are as follows: First, “Residual Correlation” significantly predicts fund performance, even after controlling for fund characteristics and other active management measures. On average, funds in highest quintile portfolio as sorted by “Residual Correlation” have 0.14% higher return per month as measured by four factor alpha (alpha obtained from the Fama French three factor model augmented with the Carhart momentum factor), compared to funds in the lowest quintile portfolio. This transfer to an annual outperformance of about 1.68% per year on a risk adjusted basis. The other components of idiosyncratic volatility, “Perfect Correlation Idiosyncratic Volatility” and “Zero Correlation Idiosyncratic Volatility” do not predict fund performance. Second, “Residual Correlation” consistently predicts performance among closet index funds, and among sector funds. As one might expect, “Residual Correlation” cannot predict abnormal returns among index funds, providing supporting evidence that “Residual Correlation” is a measure of active managment.

This chapter proceeds as follows. Section 2.2 explains empirical methodology; Section 2.3 describes data and presents summary statistics. Section 2.4 provides empirical results. Section 2.5 concludes.

2.2 Empirical methodology and hypotheses

We define a new measure of active management of mutual funds, "Residual Correlation" which is the weighted average of pairwise correlations of residual returns of assets in a fund portfolio. We construct "Residual Correlation" by (i) decomposing a fund portfolio idiosyncratic volatility into two components, the "Variance Term" which contains variances of constituent asset residual returns, and the "Covariance Term" which contains covariances of asset residual returns. This is parallel to the decomposition of variance of portfolio returns into variances and covariances of asset returns in Markowitz portfolio theory; (ii) further splitting the "Covariance Term" into the product of the "Magnitude of Covariance" which contains the asset pairwise covariances assuming perfect correlations, and the "Residual Correlation". "Residual Correlation" allows us to write the portfolio idiosynratic volatility as the sum of "Zero Correlation Idiosyncratic Volatility", which is the idiosyncratic volatility of the same portfolio assuming no residual correlations, and "Perfect Correlation Idiosyncratic Volatility", which is idiosyncratic volatility of the same portfolio assuming perfect residual correlations. These two terms are scaled by "(1-Residual Correlation)" and "Residual Correlation" respectively.

We show that the three determinant components of a fund portfolio idiosyncratic volatility ("Zero Correlation Idiosyncratic Volatility", "Perfect Correlation Idiosyncratic Volatility" and "Residual Correlation") reflect different aspects of active management and we will investigate them separately; We also compare the three components with other active management measures ("Industry Concentration Index"; "Active Share" and R^2); Finally, we distinguish "Residual Correlation" as a stronger predictor of mutual fund abnormal returns.

2.2.1 Idiosyncratic volatility of fund portfolios and its decomposition

We construct a hypothetical portfolio of assets using the holding reported by funds. This is a simple weighted portfolio rebalanced daily that intentionally omits any effect of unreported holdings or any other unobserved portfolio actions. Throughout this chapter, this hypothetical portfolio is referred to as the holdings portfolio.[7] In order to calculate idiosyncratic volatility of a fund portfolio $\sigma^2_{\epsilon^p}$, first we calculate daily returns of the portfolio as the weighted average of asset returns:

$$r^p_{t,d} = \Sigma^N_{i=1} w_i r^i_{t,d} \tag{2.1}$$

We treat the holdings portfolio as a single stock, and calculate its portfolio idiosyncratic volatility $\sigma^2_{\epsilon^p}$, using the same approach that idiosyncratic volatility is calculated for individual stocks.[8] We run portfolio daily returns on the four factor

[7]For empirical tests, we calculate idiosyncratic volatility for portfolio under both buy-and-hold and daily rebalancing assumptions, with consistent results.

[8]Campbell, Lattau, Malkiel, and Xu (2001); AHXZ (2006); Bali and Cakici (2008); and Huang, Liu, Rhee and Zhang (2010, herein after HLRZ) calculate monthly idiosyncratic risk of stocks based on the standard deviation of daily residual returns. Specifically, for each month (m), for each stock (i) that has more than 15 daily return observations in the month, they run the stock's daily returns on the four factor model. Idiosyncratic risk of stock i (σ_{ϵ^i}) is calculated by multiplying the standard deviation of daily (d) residual returns ($\sigma_{\epsilon^i_{m,d}}$) by $\sqrt{n_{i,m}}$ where $n_{i,m}$ is the number of trading days of stock i during month m. σ_{ϵ^i} can also be estimated as the squared root of sum of squared of daily residual returns $\sqrt{\Sigma^{n_{i,m}}_{d=1}({\epsilon^i_{m,d}}^2)}$. Many studies of idiosyncratic volatility use expected volatility

model (Fama French three factor model augmented with the Carhart momentum factor) to obtain portfolio daily residual returns $\epsilon^p_{m,d}$. We then multiply the variance of residual returns with the number of trading days in month m ($n_{p,m}$) to obtain $\sigma^2_{\epsilon^p}$. We will call this term idiosyncratic volatility , and define idiosyncratic risk as the standard deviation of residual returns.

By construction, the residual return of a portfolio is the weighted average of asset residual returns:

$$\epsilon^p_{m,d} = \Sigma^N_{i=1} w_i \epsilon^i_{m,d} \tag{2.2}$$

The variance of both sides of equation (2.2) produces the equation:

$$\sigma^2_{\epsilon_p} = \Sigma^N_{i=1} w_i^2 \sigma^2_{\epsilon_i} + \Sigma^N_{i=1} \Sigma^N_{j=1,j\neq i} w_i w_j \sigma_{\epsilon_i} \sigma_{\epsilon_j} \rho_{\epsilon_i,\epsilon_j} = VT + CV \tag{2.3}$$

Where w_i is the weight and σ_{ϵ^i} is the standard deviation of residual returns of asset i in the portfolio; $\rho_{\epsilon^i,\epsilon^j}$ is the correlation of residual returns of asset i and asset j, i.e. $\rho_{\epsilon^i,\epsilon^j} = correlation(\epsilon^i_{m,d}, \epsilon^j_{m,d})$; N is the number of stocks held in portfolio p. Variance of portfolio residual returns can be decomposed into two components; (i) The "Variance Term" (VT) which includes N variances of asset residual returns; and (ii) The "Covariance Term" (CT) which includes $\frac{N(N-1)}{2}$ pairwise covariances of asset residual returns. Observe that the decomposition is parallel to the standard decomposition of variance of portfolio's total returns into N variances and $\frac{N(N-1)}{2}$ covariances of asset returns as shown in Markowitz portfolio theory. The "Covariance

which is unobservable and must be estimated by sophisticated parametric models. The choice of a parametric model may be essential for volatility forecasting, yet less important for researches investigating historical data to reveal hidden information.

Term" in (2.3) can be further split into the product of two components: the first component contains covariances of asset residual returns assuming perfect residual correlations, which we call the "Magnitude of Covariance" (MC); the second component is the weighted average of residual correlations, which we call the "Residual Correlation" (RC) as follow:

$$\sigma^2_{\epsilon_p} = \Sigma_{i=1}^N w_i^2 \sigma^2_{\epsilon_i} + \Sigma_{i=1}^N \Sigma_{j=1,j\neq i}^N w_i w_j \sigma_{\epsilon_i} \sigma_{\epsilon_j} \times \frac{\Sigma_{i=1}^N \Sigma_{j=1,j\neq i}^N w_i w_j \sigma_{\epsilon_i} \sigma_{\epsilon_j} \rho_{\epsilon_i,\epsilon_j}}{\Sigma_{i=1}^N \Sigma_{j=1,j\neq i}^N w_i w_j \sigma_{\epsilon_i} \sigma_{\epsilon_j}} = VT + MC \times RC \tag{2.4}$$

The "Residual Correlation" (RC) of the fund portfolio can be written as:

$$RC = \Sigma_{i=1}^N \Sigma_{j=1,j\neq i}^N \frac{w_i w_j \sigma_{\epsilon_i} \sigma_{\epsilon_j}}{\Sigma_{i=1}^N \Sigma_{j=1,j\neq i}^N w_i w_j \sigma_{\epsilon_i} \sigma_{\epsilon_j}} \times \rho_{\epsilon_i,\epsilon_j} \tag{2.5}$$

For any pairwise residual correlation of two assets $(asset_u, asset_v)$, the weight is:

$$weight_{u,v} = \frac{w_u w_v \sigma_{\epsilon_u} \sigma_{\epsilon_v}}{\Sigma_{i=1}^N \Sigma_{j=1,j\neq i}^N w_i w_j \sigma_{\epsilon_i} \sigma_{\epsilon_j}} \tag{2.6}$$

The sum of the weights of all asset pairs equals 1. The weight of each pairwise correlation of residual returns of $asset_u$ and $asset_v$ depends on both allocation weights, and the idiosyncratic risk of the two assets. Intuitively, this weighting scheme makes sense, because the pairwise residual correlation of the two assets should have a large impact on the portfolio RC when these two assets either (i) have large allocation weights in the holdings portfolio or (ii) have large idiosyncratic risk. Consider the case where a fund manager invests large allocations on Treasury bonds and Agency bonds. These two assets may have a large residual correlation, however the weight of the pairwise correlation of these two assets in the portfolio RC should be small given that both assets have low idiosyncratic risk, and thus low contribution to the fund portfolio idiosyncratic volatility. Pairwise correlations of residual returns are in the range of -1 (perfectly negative) to 1 (perfectly positive). Let us consider $\sigma^2_{\epsilon_p}$ in the two special cases where asset residual returns are perfectly correlated, or perfectly uncorrelated.

(i) If all asset residual returns are perfectly uncorrelated ($\rho_{\epsilon^i,\epsilon^j} = 0 \forall i,j$) as we typically think in theoretical models, then portfolio RC is equal to 0. Portfolio idiosyncratic volatility can be written as:

$$\sigma^2_{\epsilon_p,\rho=0} = \Sigma_{i=1}^N w_i^2 \sigma^2_{\epsilon_i} \tag{2.7}$$

(ii) If all asset residual returns are perfectly correlated ($\rho_{\epsilon^i,\epsilon^j} = 1 \forall i,j$), such as if every asset residual return is perfectly correlated to idiosyncratic shocks, then we obtain a maximum portfolio RC of 1. Portfolio idiosyncratic volatility can be written as:

$$\sigma^2_{\epsilon_p,\rho=1} = (\Sigma_{i=1}^N w_i \sigma_{\epsilon_i})^2 \tag{2.8}$$

This can be written as the sum of the "Variance Term" and the "Magnitude of Covariance" as follow:

$$(\Sigma_{i=1}^N w_i \sigma_{\epsilon_i})^2 = \Sigma_{i=1}^N w_i^2 \sigma^2_{\epsilon_i} + \Sigma_{i=1}^N \Sigma_{j=1,j\neq i}^N w_i w_j \sigma_{\epsilon_i} \sigma_{\epsilon_j} \tag{2.9}$$

This equation allows us to write "Magnitude of Covariance" as:

$$\Sigma_{i=1}^N \Sigma_{j=1,j\neq i}^N w_i w_j \sigma_{\epsilon_i} \sigma_{\epsilon_j} = (\Sigma_{i=1}^N w_i \sigma_{\epsilon_i})^2 - \Sigma_{i=1}^N w_i^2 \sigma^2_{\epsilon_i} \tag{2.10}$$

Using (2.5), (2.7) and (2.8), we can re-write (2.4) as:

$$\sigma^2_{\epsilon_p} = \sigma^2_{\epsilon_p,\rho=0} + (\sigma^2_{\epsilon_p,\rho=1} - \sigma^2_{\epsilon_p,\rho=0}) \times RC \tag{2.11}$$

Or equivalently,

$$\sigma^2_{\epsilon_p} = \sigma^2_{\epsilon_p,\rho=0} \times (1 - RC) + \sigma^2_{\epsilon_p,\rho=1} \times RC \tag{2.12}$$

Equation (2.12) demonstrates that a portfolio idiosyncratic volatility $\sigma^2_{\epsilon_p}$ is the sum of idiosyncratic volatility of the perfectly diversified portfolio ($\sigma^2_{\epsilon_p,\rho=0}$, which we call "Zero Correlation Idiosyncratic Volatility") and idiosyncratic volatility of the perfectly undiversified portfolio ($\sigma^2_{\epsilon_p,\rho=1}$, which we call "Perfect Correlation Idiosyncratic Volatility"). These are scaled by $(1 - RC)$ and RC respectively. Therefore, "Residual Correlation" is a measure representing where $\sigma^2_{\epsilon_p}$ lies in the line connecting $\sigma^2_{\epsilon_p,\rho=0}$ and $\sigma^2_{\epsilon_p,\rho=1}$. $\sigma^2_{\epsilon_p}$ would be closer to $\sigma^2_{\epsilon_p,\rho=1}$ ($\sigma^2_{\epsilon_p,\rho=0}$) if "Residual Correlation" is large (small). $\sigma^2_{\epsilon_p}$ may be less than $\sigma^2_{\epsilon_p,\rho=0}$ if "Residual Correlation" is negative. From equation (2.11), the "Residual Correlation" can be calculated as follow:

$$RC = \frac{\sigma^2_{\epsilon_p} - \sigma^2_{\epsilon_p,\rho=0}}{\sigma^2_{\epsilon_p,\rho=1} - \sigma^2_{\epsilon_p,\rho=0}} \tag{2.13}$$

This parsimonious equation is mathematically equivalent to portfolio "Residual Correlation" in equation (2.5) while eliminating the need of estimating pairwise residual correlations, which is computationally cumbersome. Instead of calculating $\frac{N(N-1)}{2}$ pairwise residual correlations, you only need to calculate the idiosyncratic volatility of the portfolio and of its N constituent assets.

Each of the three components $\sigma^2_{\epsilon_p,\rho=0}$, $\sigma^2_{\epsilon_p,\rho=1}$, and "Residual Correlation" contributes to $\sigma^2_{\epsilon_p}$ differently. They reflect different aspects of active management, and they could have different effects upon fund performance. We will analyze these three components separately to see how each contributes to $\sigma^2_{\epsilon_p}$. We identify which aspect of active management each component might reflect, and construct hypotheses about how each should affect fund performance.

2.2.1.1 "Perfect Correlation Idiosyncratic Volatility" ($\sigma^2_{\epsilon_p,\rho=1}$)

$\sigma^2_{\epsilon_p,\rho=1}$ is squared weighted average of asset idiosyncratic risk. This component reflects funds' general preferences for the level of idiosyncratic risk. If a fund tends to hold assets with high (low) idiosyncratic risk, it will have high (low) $\sigma^2_{\epsilon_p,\rho=1}$. Falkenstein (1996) shows that mutual funds have non-linear preferences for assets with high idiosyncratic risk. These high idiosyncratic risk stocks may have greater firm-specific information that an active manager may exploit. AHXZ (2006) show that stocks with high idiosyncratic risk have low average returns. In the absence of active management, a fund that spreads its holdings to a large number of high idiosyncratic risk assets will have a high $\sigma^2_{\epsilon_p,\rho=1}$, which should result in low performance due to the negative relationship identified by AHXZ.

2.2.1.2 "Zero Correlation Idiosyncratic Volatility" ($\sigma^2_{\epsilon_p,\rho=0}$)

Funds holding high idiosyncratic risk assets will also have high $\sigma^2_{\epsilon_p,\rho=0}$. However, there is a subtle, yet important difference between $\sigma^2_{\epsilon_p,\rho=1}$ and $\sigma^2_{\epsilon_p,\rho=0}$. The asset weights (which reflect the concentration of assets in a portfolio) have a much stronger influence upon the $\sigma^2_{\epsilon_p,\rho=0}$ than they do upon the $\sigma^2_{\epsilon_p,\rho=1}$. In other word, $\sigma^2_{\epsilon_p,\rho=0}$ is much more sensitive to the weights and the number of assets in the portfolio. This component will increase a lot if the fund loads heavily on a single asset, or if it reduces number of assets held.[9] Let us consider two examples: consider fund A, and Fund B, both have preferences for high idiosyncratic risk assets but have different concentrations. For simplicity, assume every constituent asset in fund A and fund B has the same idiosyncratic risk of σ_{ϵ_i}. For both assets, we have: $\sigma^2_{\epsilon_p,\rho=1} = (\Sigma_{i=1}^N w_i\sigma)^2 = (\Sigma_{i=1}^N w_i)^2 \times \sigma^2 = \sigma^2$ (i) In the first case, assume fund A is concentrated in only a few assets: fund A holds 10 assets with equal weights, while fund B holds 100 assets with equal weights. We have $\sigma^2_{\epsilon_A,\rho=0} = \Sigma_{i=1}^{10}(\frac{1}{10}^2\sigma^2) = \frac{1}{10}\times\sigma^2$ while $\sigma^2_{\epsilon_B,\rho=0} = \Sigma_{i=1}^{100}(\frac{1}{100}^2\sigma^2) = \frac{1}{100} \times \sigma^2$ (ii) In the second case, assume fund A has has many assets but chooses to load heavily on a single asset. Assume Fund A hold 90% on asset X, and 10% on asset Y while Fund B hold 50% on asset X, and 50% on asset Y. We have $\sigma^2_{\epsilon_A,\rho=0} = \Sigma(.9^2 \times \sigma^2 + .1^2 \times \sigma^2) = .82 \times \sigma^2$ while $\sigma^2_{\epsilon_B,\rho=0} = \Sigma(.5^2 \times \sigma^2 + .5^2 \times \sigma^2) = .5 \times \sigma^2$. These examples demonstrate that given the same preferences for level of idiosyncratic risk, funds have a higher concentration on idiosyncratic risk, such as they hold fewer assets, or load more heavily in a single (or a few) asset(s) would have much higher $\sigma^2_{\epsilon_p,\rho=0}$. For this reason, we claim

[9]For easy understanding of the role of weights or concentration in determining $\sigma^2_{\epsilon_p,\rho=0}$, consider a simple case where all constituent assets have the same weight $\frac{1}{N}$, then the $\sigma^2_{\epsilon_p,\rho=0}$ is equal to $\frac{1}{N} \times (\frac{1}{N} \times \Sigma_{i=1}^N \sigma^2_{\epsilon_i})$ This term is related to the ratio of average assets' idiosyncratic volatility and the number of assets in the fund portfolio. If a fund is well diversified in the sense that it holds a large number of assets, without concentrating the fund's weight on any individual asset, then all weights are small and the $\sigma^2_{\epsilon_p,\rho=0}$ is close to zero regardless of individual assets idiosyncratic volatility. If a fund concentrates to a small number of high idiosyncratic risk assets, and/or holds a large weight on a single asset with high idiosyncratic risk, the $\sigma^2_{\epsilon_p,\rho=0}$ may be considerably large. $\sigma^2_{\epsilon_p,\rho=0}$ can be approximated by $\frac{\sigma^2_{\epsilon_p,\rho=1}}{N}$ The correlation between $\sigma^2_{\epsilon_p,\rho=0}$ and $\frac{\sigma^2_{\epsilon_p,\rho=1}}{N}$ is 93%

that $\sigma^2_{\epsilon_p,\rho=0}$ reflects a fund's concentration in idiosyncratic risk, i.e. funds allocate idiosyncratic risk to a single (or a few) individual asset(s).

Funds may seek for concentration on high idiosyncratic risk assets, i.e. they might prefer big bets on idiosyncratic risk of a single (or a few) individual assets in two cases: (i) They might have superior skills and would seek to exploit their information in high idiosyncratic risk assets. If this is the case, we would expect a positive relationship between $\sigma^2_{\epsilon_p,\rho=0}$ and fund performance. (ii) They might want to gamble to increase their chances of earning high returns. These funds should not persistently outperform. In this case, we would not expect a positive relationship between $\sigma^2_{\epsilon_p,\rho=0}$ and fund performance.

Considering both $\sigma^2_{\epsilon_p,\rho=0}$ and $\sigma^2_{\epsilon_p,\rho=1}$ together, if active management can be identified by funds making big bets on high idiosyncratic risk assets (i.e those with more firm-specific information), then $\sigma^2_{\epsilon_p,\rho=0}$ should predict performance, and $\sigma^2_{\epsilon_p,\rho=1}$ might do so too. However, we expect $\sigma^2_{\epsilon_p,\rho=0}$ to have stronger predictive power of fund performance. This is because we expect a fund to profit more from concentrating its bets on a few high idiosyncratic risk assets, than it would by scattering investments across a large number of high idiosyncratic risk assets.

2.2.1.3 "Residual Correlation"

"Residual Correlation" reflects fund managers' correlated bets on idiosyncratic risk of individual assets. We claim that active fund managers anticipate positive idiosyncratic (or residual) return events. Once they acquire information, they invest their money in assets expected to provide superior returns from those events. By exposing their portfolios to those stocks, their portfolios will show correlated residual returns that are independent of systematic risk factors.

It's important to emphasize that while "Residual Correlation" reflects correlated bets, $\sigma^2_{\epsilon_p,\rho=0}$ reflects both independent bets on idiosyncratic risk and correlated bets. However, it is more strongly influenced by the independent best on idiosyncratic risk. For example, if a fund manager expects five assets to act alike in providing superior idiosyncratic returns given an incoming event, he would allocate investment among these five assets. This action of the fund manager would increase the fund's "Residual Correlation". On the other hand, if the fund manager expects this event to have a positive impact upon only one asset, he would concentrate most or all of his investment in this single asset, which would increase $\sigma^2_{\epsilon_p,\rho=0}$ rather than "Residual Correlation".

Another difference between "Residual Correlation" and $\sigma^2_{\epsilon_p,\rho=0}$ is that while "Residual Correlation" is generally related to bets on events, $\sigma^2_{\epsilon_p,\rho=0}$ is less sensitive to event related bets. For example a fund manager may make a big investment on an asset which is underpriced according to his analysis without any expected incoming events. Following one events that simultaneously impacts several assets and react accordingly might be more effective than following one single asset by studying its fundamentals and its related events such as earning surprise pertaining to this asset in order to make a relevant investment decision. For this reason, "Residual Correlation" might better identify active management and predict fund performance, compared to $\sigma^2_{\epsilon_p,\rho=0}$.

Proposition: If active fund managers have superior skills in identifying positive idiosyncratic return events which simultaneously impact various assets, "Residual Correlation" should positively predict fund performance.

2.2.2 Total Return Correlation vs. Residual Correlation

Following Markowitz portfolio theory, we take the variance both sides of equation (2.1) and decompose the variance of portfolio returns into variances and covariances of asset returns.

$$\sigma_p^2 = \Sigma_{i=1}^N w_i^2 \sigma_i^2 + \Sigma_{i=1}^N \Sigma_{j=1,j\neq i}^N w_i w_j \sigma_i \sigma_j \rho_{i,j} \tag{2.14}$$

In the same way we decompose portfolio idiosyncratic volatility, we further decompose portfolio total volatility by splitting the "Covariance Term" into the product of the "Magnitude of Covariance" which contains the pairwise covariances of asset returns assuming perfect correlations, and the "Total Return Correlation".

$$\sigma_p^2 = \Sigma_{i=1}^N w_i^2 \sigma_i^2 + \Sigma_{i=1}^N \Sigma_{j=1,j\neq i}^N w_i w_j \sigma_i \sigma_j \times \frac{\Sigma_{i=1}^N \Sigma_{j=1,j\neq i}^N w_i w_j \sigma_i \sigma_j \rho_{i,j}}{\Sigma_{i=1}^N \Sigma_{j=1,j\neq i}^N w_i w_j \sigma_i \sigma_j} \tag{2.15}$$

"Total Return Correlation" ($TRC = \frac{\Sigma_{i=1}^N \Sigma_{j=1,j\neq i}^N w_i w_j \sigma_i \sigma_j \rho_{i,j}}{\Sigma_{i=1}^N \Sigma_{j=1,j\neq i}^N w_i w_j \sigma_i \sigma_j}$) is the weighted average of pairwise correlation of asset returns, where the weight of each pairwise correlation of asset returns depends on both allocation weights, and the total risk (or total volatility) of the two assets.

Thus, "Total Return Correlation" allows us to write the portfolio return volatility σ_p^2 as the sum of $\sigma_{p,\rho=0}^2$ (which we call "Zero Correlation Total Volatility" or the volatility of the same portfolio assuming no asset return correlations)[10] and $\sigma_{p,\rho=1}^2$

[10]Notice that the ρ here is the $\rho_{i,j}$ which is the pairwise correlation of total returns of asset i and

(which we call "Perfect Correlation Total Volatility" or the volatility of the same portfolio assuming perfect asset return correlations). These are scaled by (1-"Total Return Correlation") and "Total Return Correlation" respectively. Therefore, "Total Return Correlation" is a measure representing where σ_p^2 lies in the line connecting $\sigma_{p,\rho=0}^2$ and $\sigma_{p,\rho=1}^2$.

$$\sigma_p^2 = \sigma_{p,\rho=0}^2 \times (1 - TRC) + \sigma_{p,\rho=1}^2 \times TRC \tag{2.16}$$

In which, $\sigma_{p,\rho=0}^2 = \Sigma_{i=1}^N w_i^2 \sigma_i^2$, and $\sigma_{p,\rho=1}^2 = (\Sigma_{i=1}^N w_i \sigma_i)^2$.

The "Total Return Correlation" can be calculated as follow:

$$TRC = \frac{\sigma_p^2 - \sigma_{p,\rho=0}^2}{\sigma_{p,\rho=1}^2 - \sigma_{p,\rho=0}^2} \tag{2.17}$$

"Total Return Correlation" reflects fund managers' correlated bets on total risk of individual assets, which is more strongly driven by systematic risk factors. According to Cremers an Petajisto (2009), stock selection could be measured with residual volatility from a multifactor regression of fund return on a number of systematic factor portfolios (intended to capture all exposure to systematic risk) while factor timing could be measured with tracking error (which commonly defined as the standard deviation or variance of the market adjusted return of mutual funds). Similarly, we argue that if Residual Correlation captures stock selection of fund managers, Total Return Correlation may capture factor timing of fund managers. If

asset j. This is different from the ρ in the portfolio idiosyncratic volatility or residual correlation formulas where ρ refers to $\rho_{\epsilon_i,\epsilon_j}$ which is the pairwise correlation of residual returns of asset i and asset j.

factor timing increases (or hurts) fund performance, we would expect Total Return Correlation to positively (negatively) predict fund performance.

2.3 Data and summary statistics

Our mutual fund data come from two sources. The first is the CRSP Survivorship Bias Free Mutual Fund Database includes information funds characteristics such as fund returns, total net assets, investment objective, turnover ratio, expense ratios and other types of fees. CRSP Mutual Fund Database records different share classes of the same fund as distinct funds. The second source is CDA/Spectrum S12 mutual fund holding database. The CDA database is collected from mutual funds reports filed with SEC and from funds voluntary reports. The two databases are merged using MFLINKS file of Wharton Research Data Services (WRDS).

MFLINKS allows us to match different share classes of the same fund that are recorded as distinct funds in CRSP, into the holdings of the corresponding fund in CDA/Spectrum. We aggregate different share classes of the same fund in CRSP as follows. For age and qualitative characteristics (name, investment objective), we keep the value of the oldest share class. For TNA of the fund, we add up TNAs of its share classes. For other quantitative characteristics, we take the weighted average of other quantitative characteristics (return, expense ratio, management fee, turnover ratio), using the most recent TNA of each share class as weights.[11]

Similar to Kacperczyk, Marcin, Sialm, and Zheng (2008), Kacpercyk et al

[11] Also see Kacperczyk, Sialm, Zheng (2006)

(2005) and other papers dealing with holding data of mutual funds, we assume that funds carry the same holding from a report date until the next report date (or for 6 months, whichever sooner). Quarterly Active Share and Tracking Error data for the period 1990 to 2006 is provided by Cremers and Petajisto.[12] These measures are assumed to be the same for the last month of the quarter, and the first two months of the next quarter.

We focus on actively managed domestic equity mutual funds, eliminating balanced, bond, money market, international.[13] We exclude funds with TNA less than $5 million because inclusion of smaller funds may cause a survivorship bias problem. Since a stock holding is included only if it can be matched with the CRSP stock file

[12]Active Share for the extended period 1980 to 2010 is also provided by Petajisto. The Active Share used by Petajisto (2013) may be slightly different from the Active Share used in Cremers and Petajisto (2009) due to minor changes in the selection of bechmark index. See Petajisto (2013) for more details.

[13]Similary to Huang, Sialm, and Zhang (2011), we select funds with the following Lipper objectives: 'CA', 'CG', 'CS', 'EI', 'FS', 'G', 'GI', 'H', 'ID', 'LCCE', 'LCGE', 'LCVE', 'MC', 'MCCE', 'MCGE', 'MCVE', 'MLCE', 'MLGE', 'MLVE', 'MR', 'NR', 'S', 'SCCE', 'SCGE', 'SCVE', 'SG', 'SP', 'TK', 'TL', 'UT'. When the Lipper code is missing, we select fund with the following Strategic Insights objectives: 'AGG', 'ENV', 'FIN', 'GMC', 'GRI', 'GRO', 'HLT', 'ING', 'NTR', 'SCG', 'SEC', 'TEC', 'UTI', 'GLD', 'RLE'. When both Lipper and SI codes are missing, we select funds with the following Wiesenberger objectives: 'G', 'G-I', 'G-S', 'GCI', 'IEQ', 'ENR', 'FIN', 'GRI', 'HLT', 'LTG', 'MCG','SCG', 'TCH', 'UTL', 'GPM'. If a fund has none of these objective codes but it has a CS policy or has the percentage of common shares in the porfolio between 80%-105%, then the fund will be included. The percentage of common shares is calculated as the time-series average for each fund. We exclude all funds that contain the following words in the fund name: 'INTERNATIONAL', 'GLOBAL', 'BOND', 'BALANCED', 'MONEY MARKET'.

and its Idiosyncratic Volatility can be calculated,[14] we follow Cremers and Petajisto (2009) to include only funds that have more than 67% of value of stock holdings over the fund total net asset. We also require that funds hold ten stocks or more and that fund characteristics (TNA; Age; Turnover; Expenses) are non-missing. After all exclusions, our final sample includes 3921 actively managed equity funds.[15] Our sample period is from 1980 until 2012.

Zero Correlation Idiosyncratic Volatility $\sigma^2_{\epsilon_p,\rho=0}$ of a holdings portfolio is calculated from asset idiosyncratic volatility, using formula (2.7). Perfect Correlation Idiosyncratic Volatility $\sigma^2_{\epsilon_p,\rho=0}$ is calculated from asset idiosyncratic volatility, using formula (2.8). Monthly portfolio idiosyncratic volatility $\sigma^2_{\epsilon_p}$ is the variance of portfolio daily residual return as described in the empirical methodology.[16] Residual Correlation is calculated using formula (2.13). Each of these four variables are highly volatile, so unless otherwise stated, we take the average of values in the six recent months to have the monthly value for each of these valuable.

Table 2.1 presents summary statistics of the mutual funds in our sample. The value of "Residual Correlation" is in the range -0.07 to 0.78. The mean is 0.02 and

[14]Monthly Idiosyncratic Volatility is calculated for stocks that have more than fifteen daily return observations in a month.

[15]This sample size is large because (i) The 2013 update of MFLink adds many new links and provides links between the 98% of the domestic equity funds in CRSP and Thomson-Reuters; (ii) We do not eliminate sector funds and index funds. We also tested our result with a smaller sample in which we eliminate sector funds and index funds. The result is similar.

[16]We calculated daily portfolio return from asset daily return, assuming buy-and-hold weights. However, the results are not different when we use constant asset weights (under the daily rebalancing assumption).

the median is 0.01. This suggests that the “Residual Correlation” is close to zero for most of funds.

The mean of holdings portfolio’s weighted average of asset idiosyncratic risk is 7.32 (this is the squared Perfect Correlation Idiosyncratic Volatility. Weighted average of idiosyncratic risk is reported for an easy comparison with the weighted average of idiosyncratic risk of stocks in the market). This mean is higher than the mean of market weighted average idiosyncratic risk of 6.33 (calculated for the market portfolio in the same way we do for fund holdings portfolios). This statistics is consistent with Falkenstein (1996) who shows that *“in aggregate, funds have preference towards securities with higher idiosyncratic risk”*.

The mean of fund holdings portfolio’s idiosyncratic risk is 1.44, and the median is 1.15, which are much smaller than the mean and median of fund holdings portfolio’s weighted average idiosyncratic risk. This result is expected, given the “Residual Correlation” is very small for most of funds. Remember that the idiosyncratic risk of a holdings portfolio is equal to the weighted average idiosyncratic risk of assets held by the fund only when the fund has a perfectly positive weighted average “Residual Correlation” of 1, meaning the fund does not diversify the idiosyncratic volatility of its holdings at all.

2.4 Empirical Results

In this section, we present the empirical results. First, we establish “Residual Correlation” as a credible measure of active management by (i) looking at the correlation structure of portfolio idiosyncratic volatility $\sigma^2_{\epsilon_p}$ and its three components "Residual Correlation"; $\sigma^2_{\epsilon_p,\rho=0}$ (Zero Correlation Idiosyncratic Volatility) and $\sigma^2_{\epsilon_p,\rho=1}$ (Perfect Correlation Idiosyncratic Volatility) (ii) examining the persistence in active management as indicated by the movement of "Residual Correlation". Second, we compare weighted average "Residual Correlation" with other active management measures by (i) showing and comparing the distribution of "Residual Correlation" and other active management measures; (ii) presenting the correlation structure of active management measures; and (iii) constructing two-way independent sort to see the percentage of funds in each "Residual Correlation" quintile, given the funds are in the same quintile of other active management measures. Next, we investigate the performance predictability by one-way portfolio sorts of "Residual Correlation", $\sigma^2_{\epsilon_p}$ (portfolio idiosyncratic volatility), $\sigma^2_{\epsilon_p,\rho=0}$ (Zero Correlation Idiosyncratic Volatility) and $\sigma^2_{\epsilon_p,\rho=1}$ (Perfect Correlation Idiosyncratic Volatility); and by panel data regressions. In addition, we investigate the performance predictability of "Residual Correlation" among specific groups of funds, such as Index Funds, Closer Indexers, and Sector Funds. Finally, we examine the evolution over time of "Residual Correlation" and weighted average idiosyncratic risk of mutual funds, compared to the evolution of those measures in the market portfolio.

2.4.1 Correlation structure of portfolio idiosyncratic volatility

Table 2.2 presents the correlation structure of portfolio idiosyncratic volatility $\sigma^2_{\epsilon_p}$ and its three components: "Residual Correlation", $\sigma^2_{\epsilon_p,\rho=0}$ (Zero Correlation

Idiosyncratic Volatility) and $\sigma^2_{\epsilon_p,\rho=1}$ (Perfect Correlation Idiosyncratic Volatility). Both "Residual Correlation" and $\sigma^2_{\epsilon_p,\rho=0}$ are highly correlated to $\sigma^2_{\epsilon_p}$ with the correlations of 0.67 and 0.83 respectively while the correlation between $\sigma^2_{\epsilon_p,\rho=1}$ and $\sigma^2_{\epsilon_p}$ are smaller (0.57). This result suggests that "Residual Correlation" and $\sigma^2_{\epsilon_p,\rho=0}$ are more essential than $\sigma^2_{\epsilon_p,\rho=1}$ in determining the magnitude of holdings portfolio idiosyncratic volatility. "Residual Correlation" has low correlations to $\sigma^2_{\epsilon_p,\rho=0}$ and $\sigma^2_{\epsilon_p,\rho=1}$ (0.22 and 0.18 respectively). The correlations between the three components are not very high, implying that the three components contain different information and should be investigated separately.

2.4.2 The persistence of "Residual Correlation"

Every year in our sample period, we sort funds into 5 quintiles based on "Residual Correlation". Among the funds that have non-missing data for two continuous years, we calculate the percentage of funds in a quintile that switch to each of the five quintiles in the next year.[17] Table 2.3 shows that "Residual Correlation" exhibits pretty strong persistence: the percentage in the diagonal is always largest for each quintile. This persistence is strongest among the most active funds (quintile 5) with 71% of funds in quintile 5 still remains in the same quintile in the next year. Second strongest persistence is observed among the least active funds (quintile 1), with 62% of funds in quintile 1 remains in the same quintile in the next year. Given the mean turnover of funds is close to 1, we can say that on average, mutual funds hold a different portfolio after one year. Regardless, the "Residual Correlation" is

[17]On average, about 15% of the funds have this variable missing in the next year.

still persistent over the two non-overlapping periods. This result is consistent with the literature: strong persistence evidence is also found with Active Share and R^2 measures.

2.4.3 Distribution and comparison of "Residual Correlation" and other active management measures

Panel A of table 2.5 presents the distribution of "Residual Correlation", in comparision with that of other active management measures. The distribution of "Residual Correlation" is highly positively skewed, with "Residual Correlation" of most of funds close to zero. The mass of funds with small "Residual Correlation" likely comes from closet indexers, of which the population increased sharply in 1990s. With this increase, the percentage of asset under management with low Active Share (less than 60%) went up from 1.5% in 1980 to 44.8% in 2003. The mean "Residual Correlation" (0.02) is greater than the median (0.006), even greater than the 75th percentile (0.016), suggesting that only a small groups of funds are highly actively managed. The same story can be seen in the highly positively skewed distribution of Industry Concentration Index, and highly negatively skewed distribution of R^2. Distribution of Active Share seems to be less skewed.

Panel B of table 2.5 shows the Correlation Structure of various active management measures. Specifically, this panel shows correlations between the three well documented active management measures (Industry Concentration Index; Active Share and R^2), and correlations of these three measures with the portfolio idiosyncratic volatility $\sigma^2_{\epsilon_p}$ and its three components.

The correlations of Industry Concentration Index, Active Share and $R2$ with $\sigma^2_{\epsilon_p}$ are pretty high (0.63, 0.75 and -0.64 respectively), confirming that these three active management measures are associated with stock selection of mutual funds. Although the correlations of these three measures with $\sigma^2_{\epsilon_p}$ are similar, the correlations of these three measures with the three components of $\sigma^2_{\epsilon_p}$ vary, suggesting that each of these three measures may be associated with different aspects of active management. The correlations of Industry Concentration Index with Residual Correlation, $\sigma^2_{\epsilon_p,\rho=0}$ and $\sigma^2_{\epsilon_p,\rho=1}$ are 0.49, 0.53 and 0.24. This suggests that Industry Concentration Index is associated to both independent bets (captured by $\sigma^2_{\epsilon_p,\rho=0}$) and correlated bets (captured by Residual Correlation) on idiosyncratic risk. A similar story is observed for R^2: the correlations of R^2 with Residual Correlation, $\sigma^2_{\epsilon_p,\rho=0}$ and $\sigma^2_{\epsilon_p,\rho=1}$ are -0.44, - 0.57 and -0.21. Active Share, however, has a very strong correlation with $\sigma^2_{\epsilon_p,\rho=0}$ (0.74) and much smaller correlation with Residual Correlation (0.37). This suggests that Active Share is associated more with independent bets on idiosyncratic risk of individual assets, than with correlated bets. The correlation of Active Share with $\sigma^2_{\epsilon_p,\rho=1}$ is also high (0.62), suggesting that Active Share is also associated with preferences for high idiosyncratic risk assets.

The correlations between "Residual Correlation" with Industry Concentration Index, Active Share and R^2 are 0.49, 0.37 and -0.44 respectively, lower than the correlations among these three active management measures (0.5, -0.55 and -0.55). This suggests that "Residual Correlation" is indeed a new dimension of active management which may reveal new information not previously observed by the other active management measures.

2.4.4 Performance predictability

We investigate the performance predictability of holdings portfolio idiosyncratic volatility and its three components. We also study the persistence of performance predictability of Residual Correlation which is only consistent predictor of fund performance among the three components of fund portfolio idiosyncratic volatility. In addition, we examine the performance predictability after controlling for fund characteristics and other active managements measures. Furthermore, we show portfolio evidence of performance predictability of Residual Correlation.

2.4.4.1 Performance Predictability of $\sigma^2_{\epsilon_p}$ and its three components (Residual Correlation, $\sigma^2_{\epsilon_p,\rho=0}$ and $\sigma^2_{\epsilon_p,\rho=1}$)

Panel A, table 2.5 presents the panel data regressions of the next month four factor alpha (or next quarter alpha) on holdings portfolio idiosyncratic volatility $\sigma^2_{\epsilon_p}$ and on each of its three components (Residual Correlation, $\sigma^2_{\epsilon_p,\rho=0}$ and $\sigma^2_{\epsilon_p,\rho=1}$), controlling for other fund characteristics which have been shown in literature to be related to performance. Next month alphas are computed with respect to the four factor model, using the factor loadings estimated by the regression of previous 36 months of monthly reported gross return (reported net returns plus expense ratio). Next quarter alpha are sum of the alphas in the next three months. We can see that $\sigma^2_{\epsilon_p}$ significantly predicts performance. Among the three components of $\sigma^2_{\epsilon_p}$, only Residual Correlation significantly predicts performance. $\sigma^2_{\epsilon_p,\rho=0}$ tend to predict performance, even though not significant. $\sigma^2_{\epsilon_p,\rho=1}$ does not predict fund performance. Some other fund characteristics are also related to fund performance: lagged performance[18] is positively related to performance confirming the finding about persistence in fund performance; Turnover and Fund size (TNA) are negatively related to performance. However these results are not very consistent across various models.

In panel B, table 2.5, we further test the performance predictability of Residual Correlation, using various performance measures. In the first two models, dependent variables are next month four factor alphas (before expenses and after expenses). In the next two models, dependent variables are the next quarter alphas (before and after expenses). In all four models, Residual Correlation significantly predicts future performance. The performance predictability is much stronger when next quarter alphas are used, probably because these models reduced the possible influence of the short term return reversal phenomenon. With regards to other control variables, we also find some evidence that lagged alphas have significant positive coefficients while

[18]Lagged performance is slightly different in each model. When the dependent variable is next quarter alpha (or CS (Characteristic Selectivity), CT (Characteristic Timing)), the lagged dependent variable is previous quarter alpha (or previous month CS, CT). When the dependent variable is next month alpha, we use the alpha estimated from the four factor benchmark regression of return in previous 36 months for lagged performance. We do not use previous month alpha as a lagged performance variable in this case because with the short term reversal phenomenon, we do not expect to see funds performance in a month to positively predicts funds performance in the next month

Turnover and TNA have significant negative coefficients. Expenses have significantly negative coefficients (when net return is used), meaning that higher expenses funds do not fully compensate investors with a better performance.

Using benchmarks based on the characteristics (market capitalization, book-to-market and prior-year return) of stocks held in mutual fund portfolios, Daniel, Grinblatt, Titman and Wermers (1997) construct "Characteristic Timing" (CT) and "Characteristic Selectivity" (CS) measures to detect whether portfolio managers successfully time their portfolio weightings on these characteristics and whether managers can select stocks that outperform the average stock having the same characteristics. CT captures fund performance resulted from factor timing activity and CS captures fund performance resulted from stock selection activity of fund managers. Columns 6 and 7 of panel B show that Residual Correlation significantly predicts Characteristic Selectivity component of returns, not the Characteristic Timing component. This suggests that Residual Correlation is associated with stock selection, not with factor timing.

2.4.4.2 Persistence of Performance Predictability of "Residual Correlation"

In this section, we examine how persistent is the performance predictability of "Residual Correlation". Panel A of table 2.6 shows that "Residual Correlation" significantly predicts fund performance (before expenses) up to five quarters in the future. The coefficients decrease over the time, in terms of both magnitude and significance. After the fifth quarter, "Residual Correlation" no longer predicts fund performance.[19] This result is consistent with the evidence in Brown and Goetzmann (1995), Bollen and Busses (2005). They demonstrate that mutual fund persistent performance remains in short period (generally less than one year). Panel B of table 2.6 shows a similar and slightly weaker result when performance after expenses is used.

2.4.4.3 Performance Predictability of various measures of active management

In this section, we study the performance predictability of "Residual Correlation", controlling for other active management measures. Table 2.7 presents the coefficients of panel data regression of funds' four factor alphas (in the next quarter) on various measures of active management, with other fund characteristics controlled for. We confirm that Industry Concentration Index significantly positively predict fund performance while R^2 significantly negatively predict fund performance as demonstrated in Kacpercyk et al (2005) and Amihud and Goyenko (2013). Active Share marginally positively predict performance. This result is consistent with Cremers and Petajisto (2009) who shows that Active Share significantly predicts performance when the fund benchmark adjusted performance is used, performance

[19]When we clean the sample differently, (for example keeping mutual funds with TNA more than one million, winsozing or deleting 0.5% of the observations based on Residual Correlation and other active management measures) the predictability may last for as short as 2 quarters to as long as more than 2 years.

predictability is not significant if the four factor alpha is used.

"Residual Correlation" significantly positively predicts fund performance, even after other active management measures are controlled for. Moreover, the coefficients of all other active management measures lose significance when Residual Correlation is added to the models. This result confirm our prediction that "Residual Correlation" reveals some information about active management not previously controlled with other active management measures.

2.4.4.4 Performance Predictability: Portfolio Evidence

We examine the relative performance of funds in quintiles of holdings portfolio "Residual Correlation". Every month, we sort all mutual funds into five quintile portfolios based on "Residual Correlation". For each quintile portfolio, we compute the equally weighted average next month gross return (reported net return plus expense ratio). The difference of equally weighted average returns of the top and the bottom quintile portfolios is also calculated. We then report the time-series average of these portfolio returns (and of the return difference of top and bottom quintile); the CAPM alpha; Fama French three factor alpha and the four factor alpha. The robust Newey-West (1987) t-statistics are also reported.

Panel A of table 2.8 exhibits that top quintile portfolio of funds sorted by "Residual Correlation" significantly outperforms the bottom quintile by various performance measures (gross raw return; as well as risk adjusted returns: CAPM alpha; Fama French three factor alpha and Carhart four factor alpha). On average, the outperformance of top quintile compared to bottom quintile is about 0.18% (gross raw return) and 0.14% (four factor alpha) per month. This transfers to an annual outperformance of 2.16% (gross raw return) and 1.68% (four factor alpha).[20]

Given the strong performance predictability of the "Residual Correlation", we further investigate if investors can exploit this relationship to make profit. In addition to gross returns, reported net returns (return after expenses) which are important for mutual fund investors are also tested. Panel B demonstrates that on an after-expenses-return basis, top quintile portfolio of funds sorted by weighted average "Residual Correlation" significantly outperform bottom quintile portfolio in the month following portfolio formation. It's been known that the holdings of mutual funds are usually not immediately publicly available at the report date. For this

[20]We also sort all mutual funds into five quintile portfolios based on portfolio idiosyncratic volatility $\sigma^2_{\epsilon_p}$ and its other two components: $\sigma^2_{\epsilon_p,\rho=0}$ and $\sigma^2_{\epsilon_p,\rho=1}$. Different with the panel data regression, there is no significant difference in performance of top and bottom quintile portfolios sorted by holdings portfolio idiosyncratic volatility. Funds with high $\sigma^2_{\epsilon_p}$ tend to have better performance, but not significant. Thus, only in rich panel data, we can see the significant predictability of holdings portfolio idiosyncratic volatility. The difference in performance of top and bottom quintile portfolios sorted by $\sigma^2_{\epsilon_p,\rho=1}$ is not significant and slightly negative. This result might be driven by the negative relationship between idiosyncratic risk and subsequent month return in the stock market as demonstrated by AHXZ (2006). Regardless of the result demonstrated by AHXZ (2006) that assets with high idiosyncratic volatility have abysmally low average returns, the insignificant relationship between weighted average idiosyncratic volatility and fund performance is not surprising. This suggests that funds do not randomly select high idiosyncratic risk assets. It's likely that they selectively pick up the assets they have superior information. There is no significant difference in performance of top and bottom quintile portfolios sorted by $\sigma^2_{\epsilon_p,\rho=0}$.

reason, we examine if investors can still profit from this performance predictability of "Residual Correlation" if fund holdings are available a quarter after the report date. Panel C and D show that the portfolio evidence of performance predictability remains valid, even when the performance in the fourth month after portfolio formation is used. The significance is loss in panel D, however the trend remains the same.

2.4.5 Active management identification among specific groups of funds

We have shown that "Residual Correlation" can identify active management and predict performance among the broad categories of equity mutual funds. In this section, we apply our measure to the three specific groups of funds: sectors fund, closet indexers and index funds, which have generally been omitted from other studies of active management. Unlike other existing active management measures, "Residual Correlation" does not depend on identification in industry or benchmark measures, which makes it particularly well suited to identifying active management among these more specialized groups of funds.

Panel A of table 2.9 tests active management among sector funds, defined in various ways. First, we form the group of sector funds by objective codes. Second, we define sector funds as the funds that hold only one industry in the funds portfolios. Third, sector funds are defined as the funds that hold one or two industries in the funds portfolios. Next, sector funds are defined as the funds that are in the highest decile of Industry Concentration Index. And finally, sector funds are defined as the funds that are in the highest decile of Herfindal Industry Concentration. Regardless of how we define sector funds, “Residual Correlation” consistently predicts active management performance while none of the alternative measures of active management consistently do so.

Panel B tests active management among closet indexers, which are defined as in Petajisto (2013). At the end of each month, all funds are sorted into quintiles by Active Share and tracking error. Funds in the lowest Active Share quintile (except those in the highest tracking error quintile) are defined as closet indexers. We used measures of Active Share and tracking error as used in Cremers and Petajisto (2009) for the short period of 1990 2006. In a different test, we used the measure of Active Share as used in Petajisto (2013) for the extended period of 1980-2010, with tracking error calculated relative to S&P 500 as in Wermers (2003) . In both samples, “Residual Correlation” significantly predicts active management performance while none of the other measure does so.

Panel C tests active management among index funds. Because index funds can be identified using various selection criteria, we repeat our test with several alternative index fund identification methods. Specifically, we form the group of S&P index funds using funds’ objective codes and names; the group of all other index funds (Dow Jones, Wilshire, NASDAQ and Russell) using funds’ names; as well as the broad group of all index funds using both objective codes and names. Regardless of our selection criteria, residual correlation, as well as Industry Concentration Index and R^2 failed to identify any significance of predictable returns attributable to active management. Surprisingly, active share significantly predicts negative performance in the group of Dow Jones, Wilshire, NASDAQ and Russell funds. This suggests that in this group of funds, as index funds holdings deviate from their index objectives, negative returns result. Such deviations within these index funds' holdings do not appear to be attributable to active management.

2.4.6 Active management over time

In this section, we examine the evolution over time of active management as measured by weighted average "Residual Correlation". Figure 2.1 presents time series of the means of "Residual Correlation" of fund quintiles sorted by "Residual Correlation". In addition, we also add the time series of the means of "Residual Correlation" of all funds and the time series of the "Residual Correlation" of the broad market portfolio.

The "Residual Correlation" of the broad market portfolio is stable, and persistently close to zero over time, while the mean of "Residual Correlation" of all mutual funds has been sharply increasing over the time. This increase is mainly driven by the tremendous rise of "Residual Correlation" of the funds in the highest quintile. The mean of "Residual Correlation" slightly increases for funds in quintile 4 and quintile 3, and remains stable for funds in the lowest two quintiles. We can say that funds that are actively managed have become more intensively managed over the time. Cremers and Petajisto (2009) present a "clear time trend toward lower active share" by showing the decreasing percentage of the actively managed funds over the time. Their result is driven by the rise in the number and the size of closet indexers since 1990s and does not reveal that the actively managed funds become more intensively managed over the time.

One may argue that the increasing trend of active management in the actively managed funds is related to the evolution of idiosyncratic volatility (or Idiosyncratic risk) in the market, given there is a small but significantly positive correlation between idiosyncratic volatility and "Residual Correlation". We will show that this is not the case, by presenting the evolution of idiosyncratic risk over time.

Figure 2.1 presents the time-series of the mean of weighted average idiosyncratic risk of all funds, as well as of funds in quintile portfolios sorted by weighted average idiosyncratic risk. The weighted average idiosyncratic risk of all stocks in the market index is also reported. We can see that the means of weighted average idiosyncratic risk of all funds, as well as of quintile portfolio of funds closely move with the weighted average idiosyncratic risk in the market. There is an increasing trend of the weighted average idiosyncratic risk of stocks in the market until late 1990s, consistent with Campbell, Lattau, Malkiel, and Xu (2001). However, after 2003 the weighted average idiosyncratic risk of stocks in the market falls back to pre-1990s levels, as demonstrated by Brandt, Brav, Graham and Kumar (2010).

Therefore, we can say that actively managed funds have become more intensively managed over time. This trend is independent from the small and stable "Residual Correlation" in the market, and also independent from the movement of idiosyncratic risk of stocks in the market.

2.4.7 Performance Predictability of Total Return Correlation

Table 2.10 test the performance predictability of Total Return Correlation, using various performance measures. In the first two models, dependent variables are next month four factor alphas (before expenses and after expenses). In the next two models, dependent variables are the next quarter alphas (before and after expenses). In all four models, Total Return Correlation does not predict future performance. The last two models show that Total Return Correlation does not predict Characteristic Selectivity component of returns, and significantly negatively predicts Characteristic Timing component of returns. This suggests that factor timing hurts, instead of increases performance of mutual funds. This result is consistent with Cremers and Petajisto (2009) who show that going from a low to high tracking error may even hurt performance.

2.4.8 Determinants of Residual Correlation

Table 2.11 show coefficients of a panel data regression on Residual Correlation. The Residual Correlation is the 12 month average of monthly Residual Correlation. And all independent variables are measured at the end of the previous year (so that it's non-overlapping with the estimation period of Residual Correlation). We use expenses, turnover, fund size, fund age and number of stocks held as explanatory variables. Because Residual Correlation and many of the independent variables are persistent over time, we cluster standard errors by funds.

We find that turnover is positively related to Residual Correlation. Fund size is also positively related to Residual Correlation, although this relationship is non-linear. Number of stocks held is negatively related to Residual Correlation. These relationships are statistically significant, yet not economically significant because all coefficients are very small. The expense ratio is not statistically significant, suggesting that more correlated bets on idiosyncratic risk of individual assets is not associated with higher expense ratio.

2.5 Conclusion

In this chapter, we propose a unifying measure that captures the asset return correlations independent of systematic risk factors. In particular, we decompose holdings portfolio idiosyncratic volatility into three components: "Zero Correlation Idiosyncratic Volatility", "Perfect Correlation Idiosyncratic Volatility" and "Residual Correlation". We show "Residual Correlation" significantly predicts fund performance, even after controlling for fund characteristics and other active management measures. On average, funds in highest quintile portfolio as sorted by "Residual Correlation" have 0.14% higher return per month as measured by four factor alpha, compared to funds in the lowest quintile portfolio. This transfer to an annual out-performance of about 1.68% per year on a risk adjusted basis. "Perfect Correlation Idiosyncratic Volatility" and "Zero Correlation Idiosyncratic Volatility" do not significantly predict fund performance. Using "Residual Correlation", we find evidence that sector funds and closet index funds are actively managed and have predictable abnormal returns.

Table 2.1: Summary Statistics

This table reports summary statistics of mutual funds in our sample. Fund characteristics and monthly net return are reported by CRSP. Gross return is calculated as reported net return add back expense ratio. Industry Concentration Index is defined as in Kacpercyk et al (2005) $ICI_t = \Sigma_{j=1}^{10}(w_{j,t} - \overline{w}_{j,t})^2$ where $w_{j,t}$ is the weight the fund loads on industry j, and $\overline{w}_{j,t}$ is the weight of industry j in the stock market. Active Share is provided by Cremers and Petajisto (2009) and is defined as $AS = \frac{1}{2}\Sigma_{i=1}^{N}|w_{fund,i} - w_{index,i}|$ where $w_{fund,i}$ and $w_{index,i}$ are the weights of asset i in the holdings portfolio and in the index, and the sum is taken over the universe of all assets. R^2 is defined as in Amihud and Goyenko (2013), obtained from regressions of fund monthly gross return (reported net return plus expense ratio) over previous 24 months on the Four Factor Model. A holdings portfolio's Residual Correlation is calculated as in equation (2.13) $RC = \frac{\sigma^2_{\epsilon_p} - \sigma^2_{\epsilon_p,\rho=0}}{\sigma^2_{\epsilon_p,\rho=1} - \sigma^2_{\epsilon_p,\rho=0}}$. Weighted Average Idiosyncratic Risk is calculated as $\sigma_{\epsilon_p,\rho=1} = \Sigma_{i=1}^{N} w_i \sigma_{\epsilon_i}$ where w_i is the weight of asset i, and the sum is taken over the universe of all assets. Monthly holdings portfolio Idiosyncratic Risk is the standard deviation of residual returns of portfolio in the four factor regression as described in the Empirical Methodology. Sample period is 1980 - 2012.

Total Number of Funds: 3,921

Variable	Mean	Median	Minimum	Maximum
Number Of Assets	149	74	10	3675
TNA in millions	1165.33	201.40	5.00	202305.80
Age (Years)	13.62	9.44	0.19	88.19
Expenses (%)	1.2118	1.1805	0.0035	14.0500
Turnover (%)	91.4177	63.0000	0.0000	6070.0000
Gross return (%)	0.77	1.21	-49.84	149.33
Net return (%)	0.67	1.12	-49.93	149.31
Industry Concentration Index	0.1024	0.0362	0.0002	1.1174
Active Share	0.7296	0.7892	0.0002	0.9993
R^2	0.8921	0.9291	0.0113	0.9999
Residual Correlation	0.02	0.01	-0.07	0.78
Weighted Average Idiosyncratic Risk	7.32	6.61	0.62	37.40
Market Weighted Average Idiosyncratic Risk	6.33	6.01	3.74	14.46
holdings portfolio Idiosyncratic Risk	1.44	1.15	0.17	11.04

Table 2.2: Correlation Structure of holdings portfolio Idiosyncratic Volatility and its three components

This table shows the time-series average of cross-sectional Spearman correlations between the holdings portfolio idiosyncratic volatility $\sigma^2_{\epsilon_p}$ and its three components: Residual Correlation; Zero Correlation Idiosyncratic Volatility $\sigma^2_{\epsilon_p,\rho=0}$ and Perfect Correlation Idiosyncratic Volatility $\sigma^2_{\epsilon_p,\rho=1}$. Monthly holdings portfolio idiosyncratic volatility is the variance of daily residual return of portfolio in the four factor regression as described in the Empirical Methodology. A fund portfolio Residual Correlation, $\sigma^2_{\epsilon_p,\rho=0}$ and $\sigma^2_{\epsilon_p,\rho=1}$ are defined as in equation (2.13), (2.7) and (2.8). Sample period is 1980 to 2012.

Variable	$\sigma^2_{\epsilon_p}$	Residual Correlation	$\sigma^2_{\epsilon_p,\rho=0}$	$\sigma^2_{\epsilon_p,\rho=1}$
$\sigma^2_{\epsilon_p}$	1			
Residual Correlation	0.67***	1		
$\sigma^2_{\epsilon_p,\rho=0}$	0.83***	0.22***	1	
$\sigma^2_{\epsilon_p,\rho=1}$	0.57***	0.18***	0.60***	1

Table 2.3: Residual Correlation Persistence

This table shows the persistence of fund portfolios' Residual Correlation. A fund portfolio's Residual Correlation is defined as in equation (2.13) $RC = \frac{\sigma^2_{\epsilon_p} - \sigma^2_{\epsilon_p,\rho=0}}{\sigma^2_{\epsilon_p,\rho=1} - \sigma^2_{\epsilon_p,\rho=0}}$. Every year in the sample period (1980 - 2012), funds are sorted into quintiles of 12-month-average of Residual Correlation. The proportion of funds switching to each quintile of next year 12-month-average of Residual Correlation is reported.

	Quintile sorted by next year 12-month-average of Residual Correlation					
Quintile sorted by 12-month-average of Residual Correlation	1	2	3	4	5	Total
1	61.55%	20.59%	9.48%	6.33%	2.04%	100%
2	21.15%	39.18%	23.94%	12.44%	3.29%	100%
3	9.13%	24.79%	35.80%	23.61%	6.68%	100%
4	5.72%	12.31%	24.46%	41.51%	15.99%	100%
5	2.46%	3.70%	6.33%	16.54%	70.97%	100%

Table 2.4: Distribution and comparison of Residual Correlation with other active management measures
Panel A presents the Min, Minimum, Maximum and various Percentiles (5th, 10th, 25th, 50%, 75th, 90th and 95th) of various active management measures, including Residual Correlation; Industry Concentration Index; Active Share and R^2. A fund portfolio's Residual Correlation is calculated as in equation (2.13) $RC = \frac{\sigma^2_{\epsilon_p} - \sigma^2_{\epsilon_p,\rho=0}}{\sigma^2_{\epsilon_p,\rho=1} - \sigma^2_{\epsilon_p,\rho=0}}$. Industry Concentration Index is defined as in Kacpercyk et al (2005) $ICI_t = \Sigma_{j=1}^{10}(w_{j,t} - \overline{w}_{j,t})^2$ where $w_{j,t}$ is the weight the fund loads on industry j, and $\overline{w}_{j,t}$ is the weight of industry j in the stock market. Active Share is provided by Cremers and Petajisto (2009) and is defined as $AS = \frac{1}{2}\Sigma_{i=1}^{N}|w_{fund,i} - w_{index,i}|$ where $w_{fund,i}$ and $w_{index,i}$ are the weights of asset i in the holdings portfolio and in the index, and the sum is taken over the universe of all assets. R^2 is defined as in Amihud and Goyenko (2013), obtained from regressions of fund monthly gross return (reported net return plus expense ratio) over previous 24 months on the Four Factor Model.Sample period is 1980 - 2012.

Panel B shows the time-series average of cross-sectional Spearman correlations between various measures of active management.

Panel A: Distribution of various active management measures

Variable	Mean	Minimum	5th Pctl	10th Pctl	25th Pctl	50th Pctl	75th Pctl	90th Pctl	95th Pctl	Maximum
Residual Correlation	0.0227	-0.0735	-0.0064	-0.0045	0.0002	0.0062	0.0161	0.0507	0.1080	0.7775
Industry Concentration Index	0.1024	0.0002	0.0057	0.0091	0.0176	0.0362	0.0774	0.3645	0.5424	1.1174
Active Share	0.7296	0.0002	0.2862	0.4750	0.6363	0.7892	0.8871	0.9249	0.9376	0.9993
R^2	0.8921	0.0113	0.6442	0.7646	0.8703	0.9291	0.9637	0.9826	0.9899	0.9999

Panel B: Correlation Structure

Variable	Industry Concentration Index	Active Share	R^2
Industry Concentration Index	1		
Active Share	0.50***	1	
R^2	-0.55***	-0.55***	1
Residual Correlation	0.49***	0.37***	-0.44***
$\sigma^2_{\epsilon_p}$	0.63***	0.75***	-0.64***
$\sigma^2_{\epsilon_p,\rho=0}$	0.53***	0.74***	-0.57***
$\sigma^2_{\epsilon_p,\rho=1}$	0.24***	0.62***	-0.21***

Table 2.5: Performance Predictability of $\sigma^2_{\epsilon_p}$ and its three components (RC, $\sigma^2_{\epsilon_p,\rho=0}$ and $\sigma^2_{\epsilon_p,\rho=1}$)
Panel A presents the coefficients of panel data regressions over the sample period from 1980 to 2012. The t-statistics (in parentheses) are based on robust standard errors (White (1980)), clustered by time and by fund. Dependent variables are next month (or next quarter) alphas computed with respect to the four factor model, using the factor loadings estimated by the regression of previous 36 months of monthly reported gross return (reported net return plus expense ratio, for before expense performance). *** 1% significance, ** 5% significance, * 10% significance.
Panel B extends the Performance Predictability of Residual Correlation, using various performance measures: next month (or next quarter) alphas (both before and after expenses) and two DGTW (1997) measures of performance (Characteristic Selectivity (CS) and Characteristic Timing (CT), which capture fund performance resulted from factor timing activity and fund performance resulted from stock selection activity of fund managers).

Panel A: Performance Predictability of Portfolio Idiosyncratic Volatility and ts three components

Explanatory variables	Next Month Alpha (Before Expenses)				Next Quarter Alpha (Before Expenses)			
$\sigma^2_{\epsilon_p}$	0.0118*				0.0379***			
	(1.7939)				(2.6955)			
Residual Correlation		1.0677*				3.749***		
		(1.8893)				(3.5786)		
$\sigma^2_{\epsilon_p,\rho=0}$			0.0187				0.0537	
			(1.1866)				(1.4302)	
$\sigma^2_{\epsilon_p,\rho=1}$				0.0012				0.0031
				(0.8085)				(1.2504)
Dependent Variable (lagged)	0.1237*	0.1356*	0.1497**	0.1449**	0.0345	0.036	0.0369	0.037
	(1.7149)	(1.8529)	(2.0331)	(2.0407)	(1.2185)	(1.2539)	(1.282)	(1.2901)
Expenses	-0.0271	-0.011	-0.028	-0.0286	-0.0818	-0.0281	-0.0736	-0.0707
	(-1.0495)	(-0.4263)	(-1.1711)	(-1.0459)	(-1.5522)	(-0.5605)	(-1.5351)	(-1.3644)
Turnover	-0.0002	-0.0001	-0.0001	-0.0001	-0.0006*	-0.0005*	-0.0004	-0.0005*
	(-1.1245)	(-1.0088)	(-0.7622)	(-1.1699)	(-1.8531)	(-1.8224)	(-1.4346)	(-1.8733)
$Log(age)$	0.0082	0.0087	0.0094	0.0176	0.015	0.0162	0.0148	0.0365
	(0.5523)	(0.5808)	(0.6314)	(1.2824)	(0.4789)	(0.5052)	(0.4639)	(1.1798)
$Log(TNA)$	-0.0223	-0.0288*	-0.015	-0.0242	-0.0455	-0.0671*	-0.0173	-0.0443
	(-1.3602)	(-1.7161)	(-0.8507)	(-1.4311)	(-1.1427)	(-1.6646)	(-0.398)	(-1.0864)
$Log(TNA)^2$	0.0006	0.0011	-0.00004	0.0005	0.0003	0.002	-0.0018	-0.0002
	(0.466)	(0.8369)	(-0.0343)	(0.3914)	(0.1065)	(0.6375)	(-0.5407)	(-0.0617)

Panel B: Performance Predictability of Residual Correlation

	Dependent variable					
Explanatory variables	Next Month Alpha (before expenses)	Next Month Alpha (after expenses)	Next Quarter Alpha (before expenses)	Next Quarter Alpha (after expenses)	CS	CT
Residual Correlation	1.0677*	1.0659*	3.749***	3.7431***	1.0173*	0.1888
	(1.8893)	(1.8859)	(3.5786)	(3.5737)	(1.862)	(-0.62)
Dependent Variable (lagged)	0.1356*	0.1363*	0.036	0.036	0.0334	0.0651
	(1.8529)	(1.8535)	(1.2539)	(1.2538)	(0.9806)	(0.83)
Expenses	-0.011	-0.0827***	-0.0281	-0.2639***	-0.0008	0.0158*
	(-0.4263)	(-3.0763)	(-0.5605)	(-5.2064)	(-0.0289)	(-1.81)
Turnover	-0.0001	-0.0001	-0.0005*	-0.0005*	-0.00003	0.0001**
	(-1.0088)	(-1.0118)	(-1.8224)	(-1.8388)	(-0.2495)	(-2.01)
$Log(age)$	0.0087	0.0087	0.0162	0.0151	0.0245	0.026
	(0.5808)	(0.5843)	(0.5052)	(0.4698)	(1.4516)	(0.48)
$Log(TNA)$	-0.0288*	-0.0279*	-0.0671*	-0.0648	-0.0123	0.0133
	(-1.7161)	(-1.6601)	(-1.6646)	(-1.6086)	(-0.6969)	(-1.31)
$Log(TNA)^2$	0.0011	0.001	0.002	0.0019	0.0002	0.0013
	(0.8369)	(0.7828)	(0.6375)	(0.5957)	(0.1935)	(0.52)

Table 2.6: Persistence of Performance Predictability

This table presents the coefficients of panel data regressions over the sample period from 1980 to 2012. The t-statistics (in parentheses) are based on robust standard errors (White (1980)), clustered by time and by fund. Independent variables are of previous month. Dependent variable are alphas in future quarters (quarter 1, 2, 3, 4, and 5), computed with respect to the four factor model, using the factor loadings estimated by the regressions of previous 36 months of gross return (reported net return plus expense ratio, for before expense performance in panel A) or monthly reported net return (after expense performance, panel B). *** 1% significance, ** 5% significance, * 10% significance.

Panel A: Dependent Variable: Alphas of future quarters (before expenses)						
Explanatory variables	Performance in Next Quarter	Performance in Quarter 2	Performance in Quarter 3	Performance in Quarter 4	Performance in Quarter 5	Performance in Quarter 6
Residual Correlation	3.749***	3.4385***	3.1826***	2.7818***	2.3362**	1.7219
	(3.5786)	(3.4602)	(3.2798)	(2.8546)	(2.3236)	(1.6113)
Other variables (see Table 2.5)	Yes	Yes	Yes	Yes	Yes	Yes

Panel B: Dependent Variable: Alphas of future quarters (after expenses)						
Explanatory variables	Performance in Next Quarter	Performance in Quarter 2	Performance in Quarter 3	Performance in Quarter 4	Performance in Quarter 5	Performance in Quarter 6
Residual Correlation	3.7431***	3.4338***	3.1792***	2.7788***	2.3351**	1.7223
	(3.5737)	(3.456)	(3.2766)	(2.8515)	(2.3224)	(1.6117)
Other variables (see Table 2.5)	Yes	Yes	Yes	Yes	Yes	Yes

Table 2.7: Performance Predictability of various measures of active management
This table presents the coefficients of panel data regressions over the sample period from 1980 to 2012. The t-statistics (in parentheses)are based on robust standard errors (White (1980)), clustered by time and by fund. Dependent variable are alphas (after expenses) in next quarter, computed with respect to the four factor model, using the factor loadings estimated by the regression of previous 36 months of reported net return. *** 1% significance, ** 5% significance, * 10% significance.

Explanatory Variables	Dependent Variable: Next Quarter Alpha (after expense)						
Residual Correlation	3.749***				4.0959***	4.4828***	3.7827***
	(3.5786)				(2.7079)	(2.6064)	(3.0035)
Industry Concentration Index		0.795***			-0.2001		
		(3.7016)			(-0.552)		
Active Share			0.2733			0.1	
			(1.5532)			(0.6336)	
R^2				-1.1655***			0.0316
				(-2.5791)			(0.0589)
Other variables (see Table 2.5)	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 2.8: Performance of Quintile Portfolios of Funds sorted by Residual Correlation
This table presents equally weighted monthly gross returns (or net returns) and risk adjusted returns (CAPM alpha; Fama French (1993) three factor alpha; four factor alpha) for quintile portfolio of funds, sorted by Residual Correlation. Portfolios are rebalanced monthly. The table also reports the difference in the returns and risk adjusted returns between the top and bottom quintiles. The performance reported is in the next month (Panel A, Panel B) and in the four month (Panel C, Panel D) since portfolio formation. A fund portfolio Residual Correlation is calculated as in equation (2.13) $RC = \frac{\sigma^2_{\epsilon_p} - \sigma^2_{\epsilon_p,\rho=0}}{\sigma^2_{\epsilon_p,\rho=1} - \sigma^2_{\epsilon_p,\rho=0}}$. Robust Newey-West (1987) t-statistics are reported in parentheses. *** 1% significance, ** 5% significance, * 10% significance. Sample period is 1980 to 2012.

	Panel A: Performance (before expenses) in the next month					
Variable	1 Low	2	3	4	5 High	5-1
Average Gross Ret	0.91***	1.00***	1.03***	1.07***	1.09***	0.18**
	(3.69)	(3.89)	(3.80)	(3.83)	(3.82)	(2.25)
CAPM Alpha	-0.05*	0.02	0.03	0.06	0.08	0.13*
	(-1.90)	(0.50)	(0.48)	(0.88)	(1.03)	(1.73)
Fama French Three Factor Apha	-0.04*	-0.01	-0.01	0.04	0.09	0.13**
	(-1.86)	(-0.31)	(-0.30)	(0.88)	(1.28)	(1.98)
Carhart Four Factor Alpha	-0.03	-0.02	-0.03	0.04	0.11	0.14**
	(-1.12)	(-0.40)	(-0.59)	(0.71)	(1.52)	(2.04)

	Panel B: Performance (after expenses) in the next month					
Variable	1 Low	2	3	4	5 High	5-1
Average Net Ret	0.83***	0.91***	0.93***	0.97***	0.99***	0.16**
	(3.34)	(3.53)	(3.44)	(3.47)	(3.46)	(2.01)
CAPM Alpha	-0.13***	-0.07	-0.07	-0.04	-0.02	0.11
	(-5.26)	(-1.53)	(-1.13)	(-0.65)	(-0.32)	(1.46)
Fama French Three Factor Apha	-0.13***	-0.10***	-0.11***	-0.06	-0.02	0.11*
	(-5.60)	(-2.95)	(-2.68)	(-1.13)	(-0.27)	(1.68)
Carhart Four Factor Alpha	-0.11***	-0.11***	-0.12***	-0.06	0	0.12*
	(-4.47)	(-2.79)	(-2.72)	(-1.07)	(0.05)	(1.75)

	Panel C: Performance (before expenses) in the fourth month					
Variable	1 Low	2	3	4	5 High	5-1
Average Gross Ret	0.88***	0.97***	1.03***	1.03***	1.05***	0.17**
	(3.57)	(3.75)	(3.83)	(3.75)	(3.77)	(2.34)
CAPM Alpha	-0.05**	0.02	0.06	0.06	0.07	0.12*
	(-2.05)	(0.49)	(1.04)	(0.91)	(0.97)	(1.78)
Fama French Three Factor Apha	-0.05**	-0.02	0.02	0.05	0.07	0.12**
	(-1.98)	(-0.57)	(0.45)	(1.00)	(1.11)	(1.97)
Carhart Four Factor Alpha	-0.03	-0.02	0.01	0.03	0.08	0.11*
	(-1.25)	(-0.57)	(0.23)	(0.61)	(1.20)	(1.82)

	Panel D: Performance (after expenses) in the fourth month					
Variable	1 Low	2	3	4	5 High	5-1
Average Net Ret	0.79***	0.88***	0.93***	0.94***	0.94***	0.15**
	(3.23)	(3.39)	(3.47)	(3.39)	(3.39)	(2.08)
CAPM Alpha	-0.14***	-0.07	-0.03	-0.04	-0.03	0.1
	(-5.34)	(-1.58)	(-0.58)	(-0.59)	(-0.46)	(1.51)
Fama French Three Factor Apha	-0.13***	-0.11***	-0.08*	-0.05	-0.03	0.1
	(-5.53)	(-3.26)	(-1.78)	(-0.97)	(-0.54)	(1.65)
Carhart Four Factor Alpha	-0.12***	-0.11***	-0.08*	-0.06	-0.03	0.09
	(-4.52)	(-2.98)	(-1.73)	(-1.12)	(-0.40)	(1.50)

Table 2.9: Active management identification among specific groups of funds

This table tests performance predictability of active management among numerous groups of funds, including sector funds; closet indexers; and index funds. Dependent variables are next quarter alphas computed with respect to the four factor model, using the factor loadings estimated by the regression of previous 36 months of monthly reported gross return (reported net return plus expense ratio). Independent variables are measures of active management of mutual fund, with the inclusion of various control variables (lag alpha, expense ratio, turnover ratio, fund age, monthly total net asset, and squared monthly total net asset.) We show only the coefficients of active management measures. Other control variables are omitted to save space. The t-statistics (in parentheses) are based on robust standard errors (White (1980)), clustered by time and by fund. The sample period is from 1980 to 2012. *** 1% significance, ** 5% significance, * 10% significance.

Panel A: Performance Predictability of Active Management among Sector Funds

Specific group of mutual funds	Active Management Measure: Residual Correlation	Active Management Measure: Industry Concentration Index	Active Management Measure: Active Share	R^2
Sector funds, selected by objective codes	4.3309*** (3.1997)	0.7032** (2.0458)	1.2116 (1.4207)	-0.8888 (-1.2647)
Sector funds, defined as funds that hold only one industry in the fund	4.6541** (2.2975)	-1.142 (-0.7716)	7.4592* (1.8218)	-0.7571 (-0.5141)
Sectors funds, defined as funds that hold one or two industry in the fund	4.8359*** (2.6886)	-0.9224 (-0.8593)	4.1039 (1.2728)	-0.6488 (-0.5314)
Sector funds, defined as funds in the highest decile of Industry Concentration Index	4.5694*** (3.3194)	0.2546 (0.4382)	3.0229** (2.3673)	-1.066 (-1.3736)
Sector funds, defined as funds in the highest decile of Herfindal Industry Concentration	4.6313*** (3.2608)	0.3676 (0.6434)	2.5237** (2.0451)	-1.165 (-1.4802)

Panel B: Performance Predictability of Active Management among Closet Indexers

Specific group of mutual funds	Active Management Measure			
	Residual Correlation	Industry Concentration Index	Active Share	R^2
Closet Indexers, as defined by Petajisto (2013)	13.4069**	0.2164	0.2895	-0.6181
AS and TE are provided by Petajisto for 1990-2006	(1.9941)	(0.1316)	(1.0995)	(-0.8182)
Closet Indexers, as defined by Petajisto (2013)	8.3311*	-0.1664	0.2117	0.0384
AS is provided by Petajisto for extended period 1980-2010 TE is calculated as in Wermers (2003)	(1.7634)	(-0.1135)	(0.9459)	(0.044)

Panel C: Performance Predictability of Active Management among Index Funds

	Active Management Measure			
Specific group of mutual funds	Residual Correlation	Industry Concentration Index	Active Share	R^2
All Index Funds, selected by names and objective codes	1.0296 (0.7477)	-0.0535 (-0.1864)	-0.1781 (-0.6248)	0.0122 (0.0174)
S&P Index Funds, selected by names and objective codes	-1.4377 (-0.5627)	-0.5548 (-0.9129)	-0.5913 (-0.969)	0.7806 (0.756)
Other Index Funds (Dow Jones, Wilshire, NASDAQ and Russell), selected by names	2.9401 (0.2089)	0.5523 (0.336)	-1.5514** (-2.1071)	-1.9086 (-1.4111)

Table 2.10: Performance Predictability of Total Return Correlation

This table presents the Performance Predictability of Total Return Correlation, using various performance measures: next month (or next quarter) alphas (before or after expenses. These alphas are computed with respect to the four factor model (using the factor loadings estimated by the regression of previous 36 months of gross return (reported net return plus expense ratio, for before expenses) or reported net return (for after expenses)), and two DGTW (1997) measures of performance (Characteristic Selectivity (CS) and Characteristic Timing (CT), which capture fund performance resulted from stock selection activity and fund performance resulted from factor timing activity of fund managers). The sample period is from 1980 to 2012. The t-statistics (in parentheses)are based on robust standard errors (White (1980)), clustered by time and by fund. *** 1% significance, ** 5% significance, * 10% significance.

Panel A: Dependent Variable: Alphas of future quarters (before expenses)						
Explanatory variables	Next Month Alpha (before expenses)	Next Month Alpha (after expenses)	Next Quarter Alpha (before expenses)	Next Quarter Alpha (after expenses)	CS	CT
Total Return Correlation	0.1224	0.1213	0.3893	0.3852	0.1378	-0.4046*
	(0.367)	(0.3636)	(0.717)	(0.7097)	(0.5534)	(-1.8412)
Other variables (see Table 2.5)	Yes	Yes	Yes	Yes	Yes	Yes

Table 2.11: Determinants of Residual Correlation

This table show the coefficients of panel data regressions over the sample period from 1980 to 2012. The t-statistics (in parentheses) are based on robust standard errors (White (1980)), clustered by fund. Dependent variable is Residual Correlation. Independent variable are fund characteristics that might have effects on Residual Correlation. *** 1% significance, ** 5% significance, * 10% significance.

Explanatory variables	Dependent variable: RC	
Expenses	0.00165	-0.00189
	(0.76)	(-0.81)
Turnover	0.00005***	0.00005***
	(3.7)	(3.72)
$Log(age)$	-0.00111	-0.00225**
	(-1.14)	(-2.27)
$Log(TNA)$	0.00755***	0.00717***
	(4.53)	(4.21)
$Log(TNA)^2$	-0.00063***	-0.00055***
	(-4.29)	(-3.64)
Number of stocks		-0.00002***
		(-8.08)

Figure 2.1: This figure presents the time-series of the means of Residual Correlation of funds in quintiles sorted by Residual Correlation. The time-series of the means of Residual Correlation of all funds is also graphed. In addition, we add the time-series of the market Residual Correlation . A fund portfolio's Residual Correlation is defined as in equation (2.13). The market portfolio's Residual Correlation is calculated in a similar way.

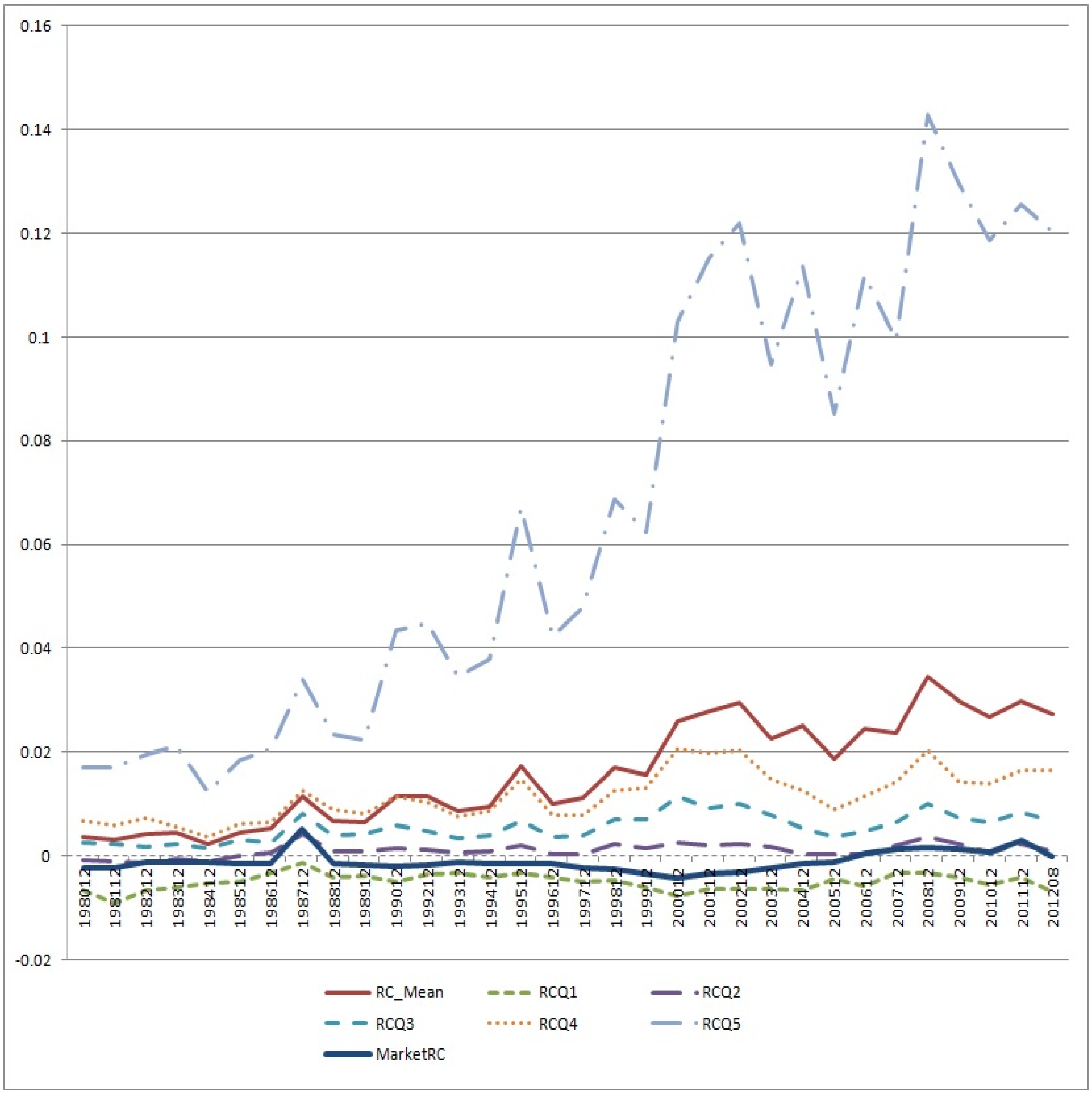
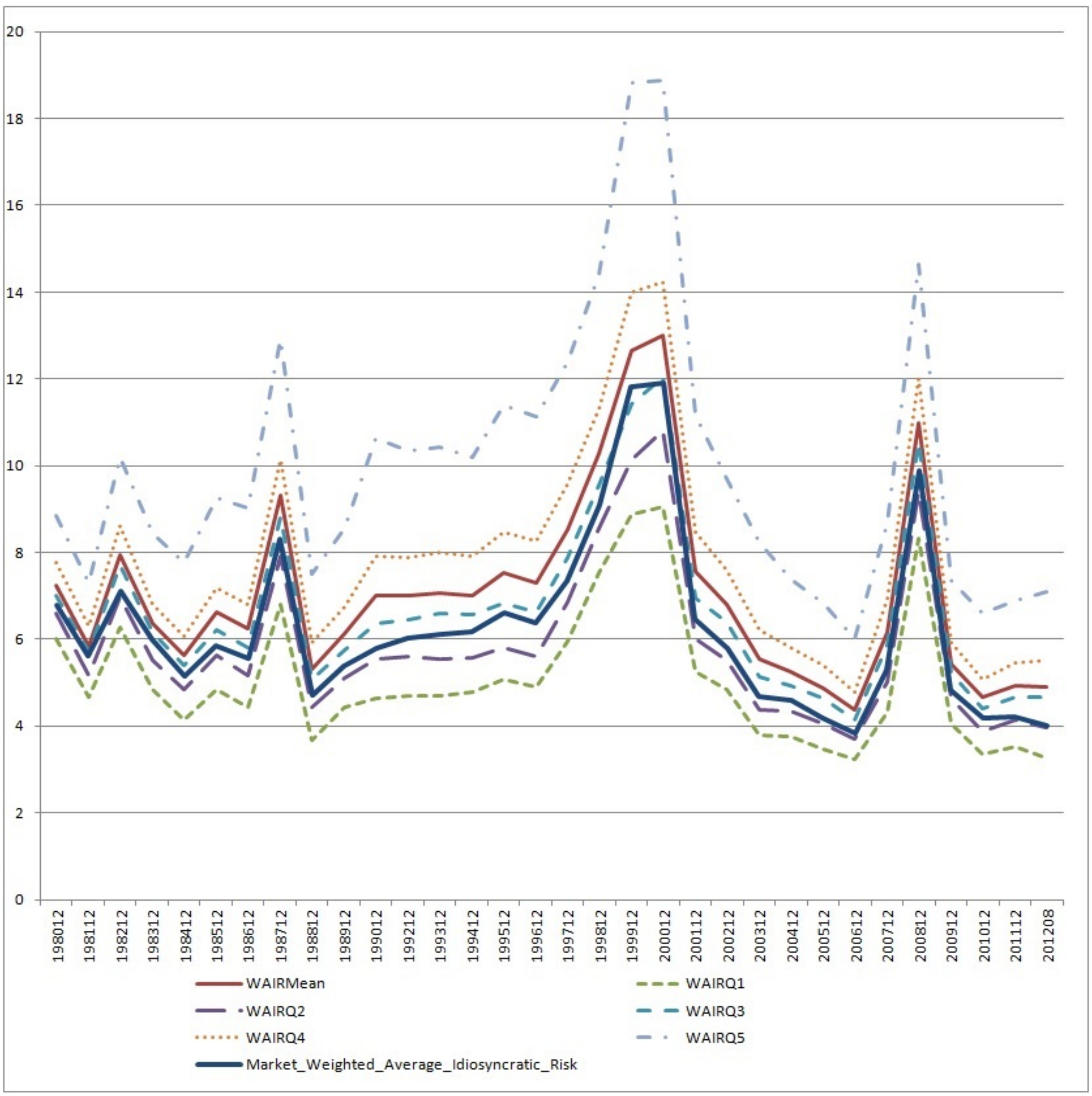


Figure 2.2: This figure presents the time-series of the means of Weighted Average Idiosyncratic Risk of funds in quintiles sorted by Weighted Average Idiosyncratic Risk. The time-series of the means of Weighted Average Idiosyncratic Risk of all funds is also graphed. In addition, we add the time-series of the market Weighted Average Idiosyncratic Risk. A fund portfolio's Weighted Average Idiosyncratic Risk is defined as $\sigma_{\epsilon_p,\rho=1} = \Sigma_{i=1}^{N} w_i \sigma_{\epsilon_i}$ where w_i is the weight of asset i in the fund portfolio, and the sum is taken over the universe of all assets held by the fund. The market portfolio's Weighted Average Idiosyncratic Risk is calculated in a similar way.



Bibliography

[1] Amihud, Yakov and Ruslan Goyenko.(2013). Mutual Fund's R^2 as Predictor of Performance. *Review of Financial Studies,* Forthcoming.

[2] Ang, Andrew, Robert J Hodrick, Yuhang Xing, and Xiaoyan Zhang. The cross-section of volatility and expected returns. *Journal of Finance, February 2006,* 61(1):259–99.

[3] Ang, Andrew, Robert J Hodrick, Yuhang Xing, and Xiaoyan Zhang. High idiosyncratic volatility and low returns: International and further U.S. evidence. *Journal of Financial Economics, January 2009*, 91(1):1–23.

[4] Bali, Turan G and Nusret Cakici. Idiosyncratic volatility and the cross section of expected returns. *Journal of Financial and Quantitative Analysis, March 2008,* 43(1):29–58.

[5] Bali, Turan G, Nusret Cakici, Xuemin (Sterling) Yan, and Zhe Zhang. Does idiosyncratic risk really matter? *Journal of Finance, April 2005,* 60(2):905–29.

[6] Bennett, J. A., Sias, R. W., and L. T. Starks. (2003).Greener Pastures and the Impact of Dynamic Institutional Preferences. *Review of Financial Studies*, 16, 1203-1238.

[7] Bennett, James and Richard Sias. Portfolio Diversification. *Journal of Investment Management,* Volume 9, number 3, 2011, pp. 74-98.

[8] Bodie, Z., Kane, A., Marcus, A. J., 2009, "*Investments,*" 8th edition, McGraw-Hill Irwin.

[9] Boehme, Rodney D, Bartley R Danielsen, Praveen Kumar, and Sorin M Sorescu. Idiosyncratic risk and the cross-section of stock returns: Merton (1987) meets miller (1977). *Journal of Financial Markets, August 2009,* 12(3):438–68.

[10] Brandt, Michael, Alon Brav, John R. Graham, and Alok Kumar, 2010, The Idiosyncratic Volatility Puzzle:Time Trend or Speculative Episodes? *Review of Financial Studies* 23, 863-899.

[11] Carhart, Mark M. On persistence in mutual fund performance. *Journal of Finance, March 1997,* 52(1):57–82.

[12] Chow, K Victor, Karen Denning, and Zhaodan Huang. Style fund investment: Performance and efficiency. *International Research Journal of Finance and Economics, November 2008,* (21):27–47.

[13] Coles, J.L., Daniel, N.D., Naveen, L. (2006).Managerial incentives and risk-taking. *Journal of Financial Economics*.79, 431–468.

[14] Constantinides, George M. Optimal stock trading with personal taxes: Implications for prices and the abnormal january returns. *Journal of Financial Economics, March 1984,* 13(1):65–89.

[15] Cram, Donald P., and D. A. Koehler. "Pollution as news: Controlling for Contempoaneous Correlation of Returns in Event Studies of Toxic Release Inventory Reporting." (2000).

[16] Cremers, K J Martijn and Antti Petajisto. How active is your fund manager? a new measure that predicts performance. *Review of Financial Studies, September 2009*, 22(9):3329–65.

[17] Cresson, John . R[superscript 2]: A market-based measure of portfolio and mutual fund diversification. *Quarterly Journal of Business and Economics, Summer-Autumn 2002*, 41(3-4):115–43.

[18] Elyasiani ,Elyas and Jingyi Jia.(2010).Distribution of institutional ownership and corporate firm performance. *Journal of Banking & Finance*, 34, 606–620.

[19] Eugene F Fama and Kenneth R French. The cross-section of expected stock returns. *Journal of Finance, June 1992*, 47(2):427–65.

[20] Eugene F Fama and Kenneth R French. Common risk factors in the returns on stock and bonds. *Journal of Financial Economics, February 1993*, 33(1):3–56.

[21] Eugene F Fama and James D MacBeth. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy, May-June 1973*, 81(3):607–36.

[22] Falkenstein, Eric G. Preferences for stock characteristics as revealed by mutual fund portfolio holdings. *Journal of Finance, March 1996*, 51(1):111–35.

[23] Fu, Fangjian. Idiosyncratic risk and the cross-section of expected stock returns. *Journal of Financial Economics, January 2009*, 91(1):24–37.

[24] Goyal, Amit and Pedro Santa-Clara. Idiosyncratic risk matters! *Journal of Finance, June 2003*, 58(3):975–1007.

[25] Huang, Jennifer, Clemens Sialm, and Hanjiang Zhang. Risk shifting and mutual fund performance. *Review of Financial Studies, August 2011*, 24(8):2575–2616.

[26] Huang, Wei, Qianqiu Liu, Ghon Rhee and Liang Zhang. (2010). Return Reversals, Idiosyncratic Risk and Expected Return. *Review of Financial Studies*, 23 (1), 147-168.

[27] Huij, Joop and Jeroen Derwall. Global equity fund performance, portfolio concentration, and the fundamental law of active management. *Journal of Banking and Finance, January 2011*, 35(1):155–65.

[28] Johnson, Timothy C. Forecast dispersion and the cross section of expected returns. *Journal of Finance, October 2004*, 59(5):1957–78.

[29] Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng. On the industry concentration of actively managed equity mutual funds. *Journal of Finance, August 2005*, 60(4):1983–2011, 2005.

[30] Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng. Unobserved actions of mutual funds. *Review of Financial Studies, November 2008*, 21(6):2379–2416, 2008.

[31] Karoui, Aymen and Iwan Meier. Performance and characteristics of mutual fund starts. *European Journal of Finance, July-September 2009*, 15(5-6):487–509.

[32] Kaushik, Abhay and Scott W Barnhart. Do mutual funds with few holdings outperform the market? *Journal of Asset Management, February 2009*, 9(6):398–408.

[33] Kosowski, Robert, Allan Timmermann, Russ Wermers, and Hal White. Can mutual fund 'stars' really pick stocks? new evidence from a bootstrap analysis. *Journal of Finance, December 2006*, 61(6):2551–95.

[34] Lehmann, Bruce N. Fads, martingales, and market efficiency. *Quarterly Journal of Economics, February 1990*, 105(1):1–28.

[35] Levy, Haim. Equilibrium in an imperfect market: A constraint on the number of securities in the portfolio. *American Economic Review, September 1978*, 68(4):643–58.

[36] Lintner, J. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics, February 1965*, 47:13–37.

[37] Miller, Edward M. Risk, uncertainty, and divergence of opinion. *Journal of Finance, September 1977*, 32(4):1151–68.

[38] Newey, Whitney K and Kenneth D West. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–08, May 1987.

[39] Petajisto, Antti. (2010). Active Share and Mutual Fund Performance. Working paper, New York University Stern.

[40] Pollet, Joshua M and Mungo Wilson. How does size affect mutual fund behavior? *Journal of Finance, December 2008*, 63(6):2941–69.

[41] Rubin, Amir and Daniel R. Smith.(2009).Institutional ownership, volatility and dividends. *Journal of Banking & Finance*, 33, 627–639.

[42] Shleifer, Andrei and Robert W Vishny. The limits of arbitrage. *Journal of Finance, March 1997*, 52(1):35–55.

[43] Yuan, Rongli, Jason Zezhong Xiao and Hong Zou.(2008).Mutual funds' ownership and firm performance: Evidence from China. *Journal of Banking & Finance*, 32 , 1552–1565.