The Elusive "Stock-Picker's Market": Dispersion and Mutual Fund Performance

Jacob Epstein

Professor Emma Rasiel, Faculty Advisor

Honors Thesis submitted in partial fulfillment of the requirements for Graduation with Distinction in Economics in Trinity College of Duke University

Duke University
Durham, North Carolina
2020

Acknowledgments

I would like to thank my faculty advisor, Dr. Emma Rasiel, for her guidance and feedback during the research process and her mentorship throughout my entire undergraduate career. I am also very grateful for all of my professors and teaching assistants in the economics department for providing a fantastic academic experience.

Abstract

This paper explores the relationship between active mutual fund performance and market dispersion from January 1990 to December 2018. I find a significant positive relationship between dispersion and 4-factor alpha overall, providing some evidence of managerial skill. There are large differences in this relationship by decade and fund selectivity. The results suggest active mutual funds were able to take advantage of stock-picking opportunities during the 1990s and 2000s, particularly the most active subset of funds. However, I find a significant negative relationship between dispersion and alpha for funds in the 2010s, indicating this relationship has changed over time. I discuss several possible explanations for this reversal, which could present interesting avenues for further research.

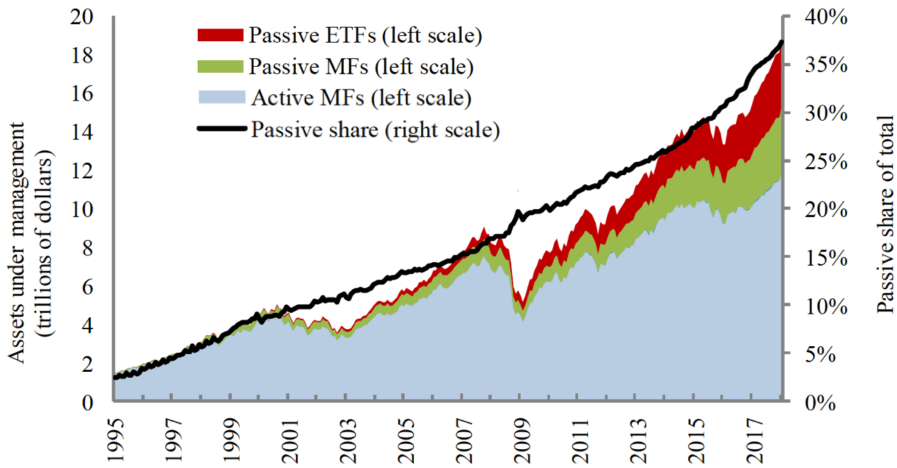
JEL classification: C30, G12, G23

Keywords: Mutual Funds, Active Management, Dispersion, Cross-Sectional Volatility

1. Introduction

In 2008, Warren Buffett famously made a $1 million bet against asset manager Ted Seides that the S&P 500 would outperform actively managed funds over a period of 10 years. By early 2017, it was clear Buffett had won; the S&P had returned 7.1% compounded annually, compared to a mere 2.2% achieved by a basket of active funds hand-picked by Seides[1]. While that contest was squarely settled, the larger debate over active versus passive management remains. Actively managed mutual funds, run by managers who try to outperform the broader market by picking stocks, have long been one of the most popular ways for retail investors to invest. In recent years, however, passive mutual funds, and other passive instruments, such as ETFs, have risen dramatically in popularity. These funds aim to mirror the returns of specific market indices and charge much lower fees than charged by active funds. A visualization of these trends is shown in figure 1. Some proponents of passive investing argue that active managers' performance is a function of luck rather than skill. In his popular 1973 book, *A Random Walk Down Wall Street*, Princeton economist Burton Malkiel claimed, "a blindfolded monkey throwing darts at a newspaper's financial pages could select a portfolio that would do just as well as one carefully selected by experts." Since then, a large body of research has explored whether active managers add value. In this paper, I evaluate whether active mutual fund managers demonstrate skill by examining whether actively managed mutual funds are able to outperform when their opportunity set is larger—periods which could be referred to as "stock-pickers' markets".

[1]https://fortune.com/2017/12/30/warren-buffett-million-dollar-bet/



Source: Morningstar, Inc.

Figure 1: Trends in active and passive management

I use dispersion, the cross-sectional volatility of the market, as a proxy for the extent to which there are stock picking opportunities. At high levels of dispersion, stock returns are very different from each other, providing ample opportunity for a stock picker to identify winners and losers. At the other extreme, when dispersion is low, all stocks perform similarly, so there is limited opportunity for stock pickers to outperform. If active managers are skilled at stock picking, one would expect to see a positive relationship between fund performance and dispersion. I first examine whether there is evidence of such a relationship. I then examine whether this relationship has changed over time. Finally, I evaluate whether a fund's performance during periods of high dispersion provides meaningful insights from the perspective of an individual considering an investment in that fund.

2. Literature Review

2.1) Performance Evaluation Framework

While there exist many frameworks for evaluating mutual fund performance, one of the most commonly used and empirically tested frameworks is the four-factor model developed by Carhart (1997) and Fama and French (1992, 1993). This approach estimates a portfolio's excess returns using the market's excess returns, as well as size, value and momentum factors. I use this framework to evaluate the performance of mutual funds in this paper.

2.2) Active vs. Passive Management

Many studies suggest that active managers tend to underperform passive managers. In one of the first studies on the topic, Jensen (1968) examined mutual fund returns from 1945-1964 and found little evidence of any individual fund outperforming the market more often than what would be expected from random chance alone. Similarly, Davis (2001) used a Fama-French three-factor model to analyze the returns of around five thousand actively managed mutual funds from 1968 – 1998 and found no evidence of positive abnormal returns. These results held when grouping by fund style and evaluating each style individually. Jones and Wermer (2011) surveyed literature on the value of active management and concluded that the average active manager does not outperform but a significant minority do add value. They also argue that comparing active funds to a particular index directly is not a fair comparison, as one cannot invest in the index directly. Instead, an investor would have to invest in a passive fund or ETF which tracks the corresponding index and charges a small fee to do so. After taking this fee into account, Jones and Wermer find that active and passively managed funds perform approximately the same.

Cremers and Petajisto (2009) propose a new measure to quantify the extent to which a mutual fund is actively managed and argue that this measure positively predicts fund performance. The new metric, “active share”, is calculated as the sum of absolute differences between weightings of stocks in a portfolio with stock weightings in the corresponding benchmark index. Frazzini, Friedman and Pomorski (2016) replicate this study but report their results separately for each benchmark index and find that this metric is as likely to correlate positively with returns as it is to correlate negatively when grouping by benchmark. Regardless, the authors note that this metric is commonly used by institutional money managers in determining fund allocations. This metric is difficult to calculate systematically because it requires knowledge of a fund’s target benchmark and weighting of each stock within that benchmark. Amihud and Govenko (2013) propose a different proxy for fund activity, $1-R^2$: the percentage variability in fund return not explained by market returns. This metric identifies fund variance due to idiosyncratic factors; funds that are more active will inherently have more idiosyncratic variance. Since this metric is easily systematically applied to many funds, I use this $1-R^2$ metric, which I refer to as selectivity, as a proxy for fund activity.

2.3) Dispersion

Several papers have explored cross-sectional volatility of stock returns, or dispersion, particularly in the context of comparison to cross-sectional volatility of fund returns. This measure is calculated as either the value-weighted or equal-weighted standard deviation of security returns in the market in a given period and quantifies the variability of security returns within a given portfolio. Ankrim and Ding (2002) and de Silva, Sapra and Thorley (2001) both comment on dispersion of securities as a driver of large variation in mutual fund returns during late 1990s and early 2000s. Petajisto (2013) explores another aspect of dispersion, using it as a proxy for attractiveness of a stock picker’s opportunity set. He argues that periods of high dispersion represent “stock-pickers’ markets,” where large differences in returns between the best and worst-performing stocks provide ample opportunity for active managers to identify winners and losers. He then finds a modest positive relationship between dispersion and excess return on a benchmark for the most active subset of funds between 1990 and 2009, supporting the idea of a stock-picker's market. When including all funds, not just the most active subset, this relationship was not significant. This paper expands on Petajisto's analysis by incorporating more funds into the analysis as well as more recent data on these funds, using a four-factor alpha instead of simple outperformance, and allowing differences in this relationship over time.

2.4) Survivorship Bias

As with any study of mutual funds, survivorship bias is an important topic to address before conducting analysis. Carhart, Carpenter, Lynch and Musto (2002) discuss at length sources of bias in studies of mutual funds. In particular, they demonstrate that mutual funds tend to disappear following poor multi-period performance, resulting in a positive bias of approximately 1% of average annual returns for studies of 15 years or more. This form of bias is unavoidable, as one cannot prevent poorly-performing mutual fund from shutting down. A more explicit form of bias is introduced when "dead" funds, or funds which are not present the end of the sample, are excluded from a sample. Many papers, such as Petajisto (2013), remark that samples which include all dead funds are essentially survivorship-bias free, as they avoid introducing this form of bias. As such, I use a sample that includes both live and dead funds in order to minimize survivorship bias.

3. Data and Methodology

3.1) Mutual Fund Data

Data on mutual fund returns come from The Center for Research in Security Prices (CRSP) Mutual Funds database. Excluding index funds, his database includes monthly data on returns and net asset values for 61,979 funds in existence for at least one month between January 1990 and December 2018. The average fund life was 103 months, resulting in a data set of 6.4 million fund-month observations. In order to exclude funds not benchmarked to equity indices, I then calculated Carhart four-factor betas for each fund over their entire lives and excluded all funds with market betas lower than 0.8 or above 1.2. I also limited the analysis to funds with lifespans of three years or greater, because at least that much data is necessary in order to calculate three-year rolling metrics. Subsetting criteria are discussed further in section 3.3. With these criteria, the final dataset includes data on 19,433 funds, with an average life of 132 months, resulting in 2.6 million fund-month observations.

3.2) Supplementary Data

Dispersion data was generated as described in section 3.4 using S&P 500 constituent data from Compustat - Capital IQ's Index Constituents database and stock returns from CRSP's Stock / Security Files database. It includes data on the cross-sectional volatility of the S&P 500, calculated monthly for a total of 348 months. Other volatility data, including monthly VIX, comes from the Wharton Research and Data Services's (WRDS) CBOE Indexes database. Fama-French factor data comes from the WRDS Fama-French Portfolio and Factors database, which in turn draws from Kenneth French's website. These data include monthly values for the risk-free rate and excess return of the market, as well as size, momentum and value factors. These are as defined by Carhart (1997) and Fama and French (1993) and reproduced below:

Risk free rate (r_f) – one month U.S. Treasury bill rate

Market return (r_{mkt}) – value-weighted returns of all NYSE, AMEX, and NASDAQ stocks

Size (r_{smb}) – difference in returns between small-cap and large-cap stocks. Calculated as the average return of three small-cap portfolios minus the average return of three large-cap portfolios.

Momentum (r_{umd}) – difference in return between high prior-period-return stocks and low prior-period-return stocks.

Value (r_{hml}) – difference in return of value stocks and growth stocks. Calculated as the average return of two value portfolios minus the average return of two growth portfolios.

3.3) Sample Identification

First, I created a subset of relevant funds using the market beta cutoffs mentioned in section 3.1. In order to estimate betas for each fund, I used a Carhart four-factor model over each fund's entire life. Its formulation shown below, where $r_{j,t}$ represents the return of fund j in period t, and $r_{mkt,t}$, $r_{smb,t}$, $r_{hml,t}$ and $r_{umd,t}$ are as described in the section 3.2:

$$r_{j,t} - r_{f,t} = \alpha_j + \beta_{mkt,j}(r_{mkt,t} - r_{f,t}) + \beta_{hml,j} r_{hml,t} + \beta_{smb,j} r_{smb,t} + \beta_{umd,j} r_{umd,t} + \epsilon_t$$

After estimating betas for each fund, I included only funds whose betas were between 0.8 and 1.2, since it is reasonable to assume these funds are benchmarked to U.S. equity indices. While this approach is somewhat arbitrary, it provides an intuitive way to exclude funds uncorrelated with U.S. equity markets, is and is easy to apply systematically as it does not require knowledge of each fund's holdings. Furthermore, the results of this paper are not sensitive to small changes in cutoff values, as can be seen in Appendix 1. A distribution of market betas for each fund is shown in figure 2. I also filtered out any funds designated as ETFs by the CRSP database, and any funds with less than 36 months of data, because it would be impossible to calculate three-year rolling metrics for these funds. By including all funds that were extant for at least three years, I avoid introducing survivorship bias.

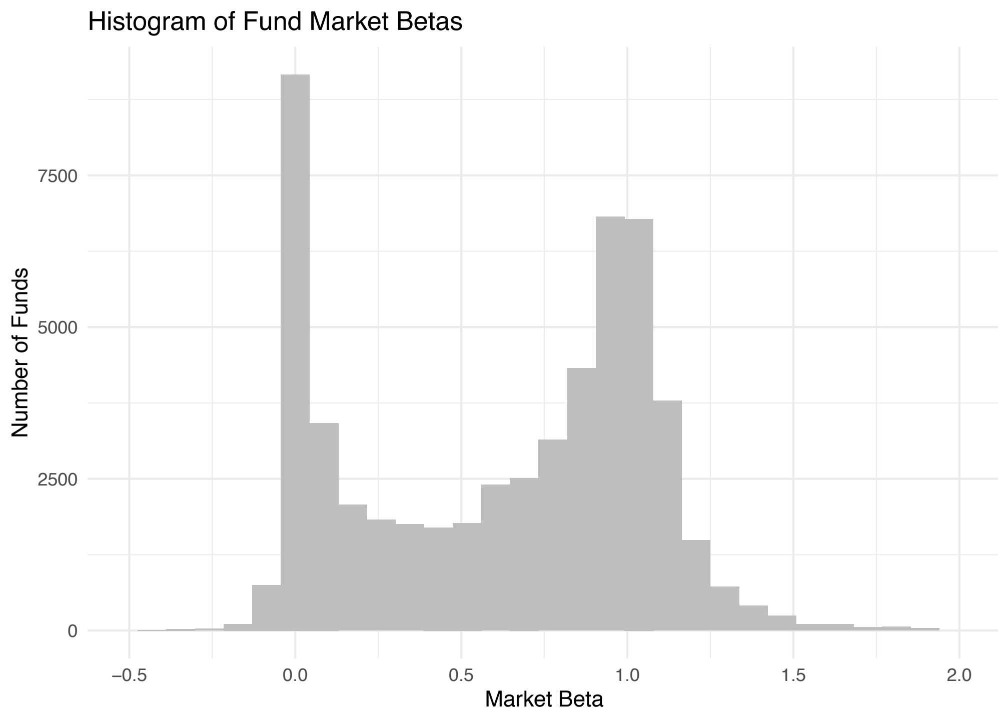


Figure 2: Distribution of market betas for all funds

3.4) Dispersion Calculation

Concurrently, I used data on S&P 500 constituents and returns to generate values for dispersion, or cross-sectional volatility of the S&P 500, equal weighted. I used S&P 500 components to calculate market dispersion rather than, say, all publicly listed equities because historical data is more readily available on these securities and these are the most commonly held stocks by mutual funds. Though Petajisto (2013) calculated dispersion on a value-weighted basis, I calculated dispersion on an equally weighted basis as it is more intuitively compatible with the idea of stock picker's markets. Large variance in stock returns present opportunities for active stock pickers regardless of the size of the stocks. Dispersion is calculated as follows, where n_t is the number of stocks in the S&P 500 in period i, $r_{j,t}$ is the return of stock j in period t and $\bar{r_t}$ is the average return of S&P 500 stocks in period i:

$$Dispersion_t = \sqrt{\frac{1}{n_t - 1}\sum_{j=1}^{n_t}(r_{j,t} - \bar{r_t})^2}$$

For context, a time series of dispersion, compared to VIX is shown below. The two are relatively strongly correlated (ρ = 0.59). Dispersion measures a different angle of market volatility than VIX; dispersion corresponds to cross-sectional variance of returns whereas VIX corresponds to time-series volatility. Historical data for both metrics are shown in figure 3.

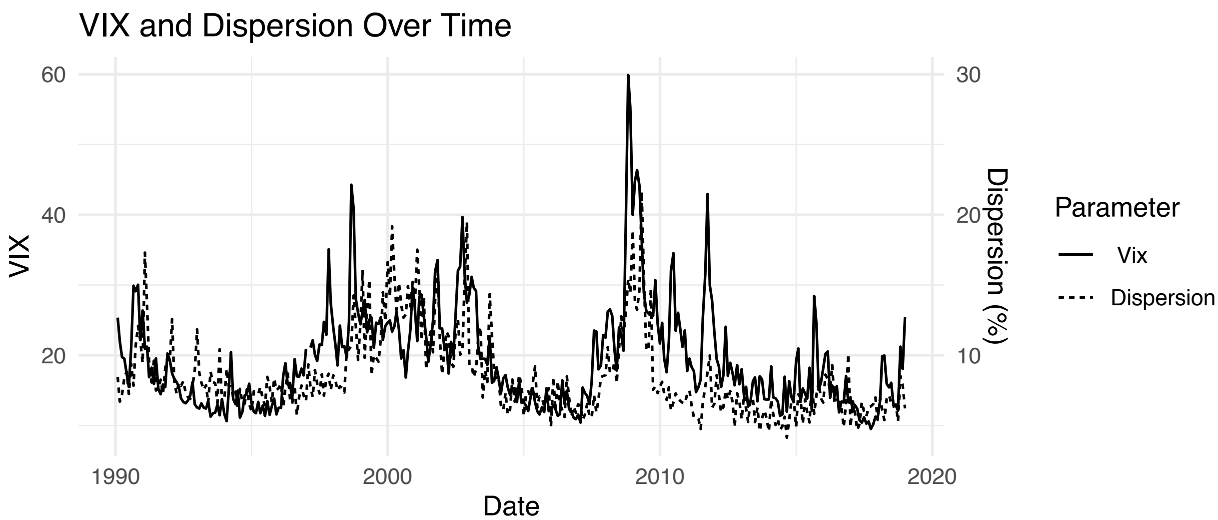


Figure 3: Historical data for VIX and Dispersion.

3.5) Fund Alpha

I then calculated Carhart four-factor alphas for each of the funds on a three-year rolling basis. For each month a fund was in the sample, I calculated its three-year rolling alpha, market beta, size beta, value beta and momentum beta using the following model:

$$r_{j,t} - r_{f,t} = \alpha_j + \beta_{mkt,j}(r_{mkt,t} - r_{f,t}) + \beta_{hml,j} r_{hml,t} + \beta_{smb,j} r_{smb,t} + \beta_{umd,j} r_{umd,t} + \epsilon_t$$

where MKT, HML, SMB and UMD are defined in section 3.2. I then calculated monthly alpha, $\alpha_{j,t}$ for each fund (not to be confused with three-year rolling alpha, α_j), defined as:

$$\alpha_{j,t} = r_{j,t} - r_{f,t} - \beta_{mkt,j}(r_{mkt,t} - r_{f,t}) - \beta_{hml,j} r_{hml,t} - \beta_{smb,j} r_{smb,t} - \beta_{umd,j} r_{umd,t}$$

Or equivalently:

$$\alpha_{j,t} = \alpha_j + \epsilon_t$$

In other words, a fund's monthly alpha is its outperformance in a given month relative to the fund's expected return given market return, risk free rates, size, value and momentum.

3.6) Selectivity

Following Amihud and Goyenko (2013), I estimated 1-R^2 using the S&P 500 as the market benchmark**.** This 1-R^2 measure can be interpreted as the percentage variability in fund return that is not explained by market returns; thus, we use this measure as a proxy for fund selectivity. First, I estimate each fund's selectivity proxy as follows:

$$1 - R_j^2 = 1 - \frac{Cov(R_j, R_{S\&P})^2}{Var(R_{S\&P})Var(R_j)}$$

I then categorize fund as high, middle or low selectivity based on this metric, both aggregated as well as for each decade. The 1990s decade is defined as January 1990-December 1999, and 2000s and 2010s are defined similarly. The motivation behind having two different selectivity labels is to allow for fund selectivity to vary over time, as one could imagine would happen as a fund shifts management or style. Low selectivity funds are defined as those in the bottom quintile of 1-R^2, high selectivity funds are those in the top quintile, and middle selectivity as those in the middle 3 quintiles. Cutoff values for each group are shown below in table 1. Percent of fund-month observations in each category are shown in table 2. 85% of observations had the same selectivity designation whether calculated overall or by decade, indicating that funds generally tend to maintain similar levels of selectivity over time.

	Low	Middle	High
Overall	(0, 0.08)	(0.08, 0.35)	(0.35, 0.99)
1990s	(0, 0.14)	(0.14, 0.54)	(0.54, 0.99)
2000s	(0, 0.08)	(0.08, 0.34)	(0.35, 0.94)
2010s	(0, 0.07)	(0.07, 0.31)	(0.31, 0.99)

Table 1: Selectivity cutoffs by decade

	Decade: High	Decade: Middle	Decade: Low
Overall: High	16%	5%	0%
Overall: Middle	4%	54%	4%
Overall: Low	0%	2%	15%

Table 2: Percent of fund-month observations in each selectivity category

3.7) Model

I then estimated three different models to understand the relationship between fund alpha and dispersion. First, I estimate the model across the entire time period, allowing for interaction with overall selectivity. VIX is also included as a covariate because, as Low (2000) points out, it represents a proxy for investor fear in the market. This model is explicitly defined below:

$$\alpha_{j,t} = \beta_0 + \beta_1 VIX_t + \beta_2 Dispersion_t + \beta_3(Middle\ Selectivity) + \beta_4(High\ Selectivity) + \beta_5(Dispersion_t : Middle\ Selectivity) + \beta_6(Dispersion_t : High\ Selectivity) + \epsilon_t$$

More simply: $\alpha_{j,t} \sim VIX_t + Selectivity_j + Dispersion_t + Selectivity_j * Dispersion_t$

Low selectivity is the default selectivity designation in this formulation and dispersion is mean centered to allow for easier interpretation of β_0. I then allowed for variation over time by adding in interaction terms for decade and used decade-by-decade measures for selectivity. The time-varying model is defined below:

$$\alpha_{j,t} = \beta_0 + \beta_1 VIX_t + \beta_2 Dispersion_t + \beta_3 2000s + \beta_4 2010s + \beta_5(Middle\ Selectivity) + \beta_6(High\ Selectivity) + \beta_7(Dispersion_t : 2000s) + \beta_8(Dispersion_t : 2010s) + \beta_9(Dispersion_t : Middle\ Selectivity) + \beta_{10}(Dispersion_t : High\ Selectivity) + \beta_{11}(Middle\ Selectivity : 2000s) + \beta_{12}(Middle\ Selectivity : 2010s) + \beta_{13}(High\ Selectivity : 2000s) + \beta_{14}(High\ Selectivity : 2010s) + \beta_{15}(Dispersion_t : Middle\ Selectivity : 2000s) + \beta_{16}(Dispersion_t : Middle\ Selectivity : 2010s) + \beta_{17}(Dispersion_t : High\ Selectivity : 2000s) + \beta_{18}(Dispersion_t : High\ Selectivity : 2010s) + \epsilon_t$$

Using the same shorthand as above, the model can be written as:

$$\alpha_{j,t} \sim VIX_t + Selectivity_j + Dispersion_t + Decade_t + Selectivity_j * Dispersion_t + Selectivity_j * Decade_t + Decade_t * Dispersion_t + Dispersion_t * Selectivity_j * Decade_t$$

I then explored dispersion in fund alpha as a function of market dispersion, comparing a simple linear relationship to a log-log relationship. Fund dispersion is defined similarly to market dispersion. Its formulation is shown below, where n_t is total number of funds in a given month, and $\bar{\alpha_t}$ is average alpha of all funds in a given month:

$$FundDispersion_t = \sqrt{\frac{1}{n_t - 1} \sum_{j=1}^{n_t} (\alpha_{j,t} - \bar{\alpha_t})^2}$$

The linear model is defined as:

$$FundDispersion_t = \beta_0 + \beta_1 Dispersion_t + \epsilon_t$$

And the log-log model is defined as:

$$log(FundDispersion_t) = \beta_0 + \beta_1 log(Dispersion_t) + \epsilon_t$$

4. Results

4.1) Initial Model

The initial model is shown below, and its corresponding output is shown in table 3:

$$\alpha_{j,t} \sim VIX_t + Selectivity_j + Dispersion_t + Selectivity_j * Dispersion_t$$

where VIX and dispersion are both mean-centered.

N	R-Squared	F	P-Value
1,865,799	0.002	674.9	0.000

	Estimate	Standard Error	T	P-Value
Intercept	(0.27)	0.04	(7.09)	0.00
VIX	(0.17)	0.00	(62.64)	0.00
Dipsersion	34.06	1.45	23.54	0.00
High Selectivity	(0.35)	0.05	(6.76)	0.00
Middle Selectivity	(0.23)	0.04	(5.32)	0.00
Dispersion x High Selectivity	3.76	1.81	2.08	0.04
Dispersion x Middle Selectivity	1.52	1.55	0.98	0.33

Table 3: Initial model coefficients

The R-squared of this regression indicates that about 0.2% of the variability in alpha can be explained by variability in the predictors. The low R-squared is likely due to the noisy nature of alphas, since these alphas represent the components of returns not already explained by general market movement, size, value and momentum. Despite this low R-squared, many of the predictors are highly significant in explaining alpha. Given the statistical significance of coefficients, we can conclude that this model provides useful information about the mean alphas for funds given selectivity and dispersion; however, the low explanatory power of this model indicates it is not particularly informative for each specific fund. A visualization of the model output is shown in figure 4, with the vertical gray bar representing mean dispersion levels.

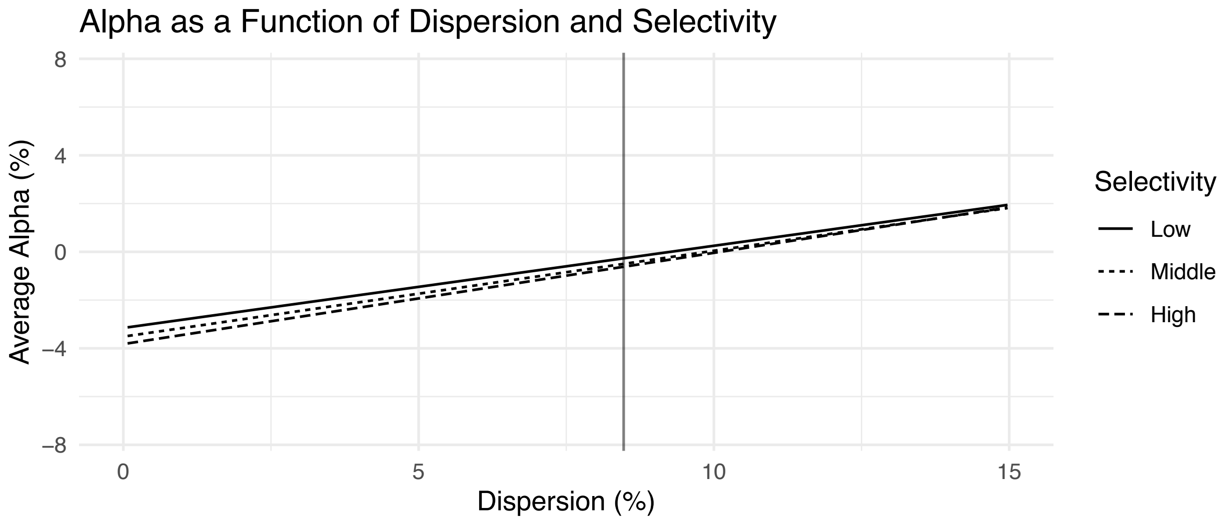


Figure 4: Initial model visualization

Low, middle and high selectivity funds had average alphas of -0.27%, -0.50% and -0.62%, all significantly different from 0 and from each other. The coefficient for VIX indicates that a 1-point increase in VIX is associate with a 17 basis point decrease in alpha. Low and middle selectivity funds had a significantly positive relationship with dispersion; a 1 percentage point increase in dispersion is associated with a 34 basis point increase in alpha for these funds. Highly selective funds have an even larger positive relationship between alpha and dispersion (38 basis points of alpha per percentage point of dispersion), suggesting that highly selective funds, are able to take advantage of high dispersion environments slightly more effectively than their less selective peers. The lower overall alpha for these funds may indicate the increased transaction costs associated with being highly active are not offset by their advantage in periods of high dispersion. This analysis implicitly assumes that fund alpha, and the relationship between dispersion and outperformance does not change over time. This assumption is likely flawed, as the investment environment has changed significantly over the past three decades. In order to account for changes in investment environment over time, the next iteration of the model allows for variation by decade.

4.2) Time-Varying Model

After adding decade as a predictor, the new model is shown below, and its output is in shown in table 4.

$$\alpha_{j,t} \sim VIX_t + Selectivity_j + Dispersion_t + Decade_t + Selectivity_j * Dispersion_t + Selectivity_j * Decade_t + Decade_t * Dispersion_t + Dispersion_t * Selectivity_j * Decade_t$$

where VIX and dispersion are again both mean-centered.

N	R-Squared	F	P-Value
1,865,799	0.008	791.9	0.000

	Estimate	Standard Error	T	P - Value
Intercept	0.07	0.15	0.47	0.64
VIX	(0.18)	0.00	(64.58)	0.00
Dispersion	(33.82)	5.47	(6.18)	0.00
2000s	(0.41)	0.17	(2.42)	0.02
2010s	(1.21)	0.17	(7.04)	0.00
Highly Selective	1.58	0.22	7.13	0.00
Middle Selective	0.37	0.18	2.09	0.04
Dispersion x 2000s	84.37	5.77	14.62	0.00
Dispersion x 2010s	19.18	6.82	2.81	0.00
Dispersion x Highly Selective	295.35	7.75	38.12	0.00
Dispersion x Middle Selective	74.77	6.28	11.91	0.00
2000s x Highly Selective	(1.85)	0.24	(7.69)	0.00
2010s x Highly Selective	(6.48)	0.25	(25.56)	0.00
2000s x Middle Selective	(0.91)	0.19	(4.67)	0.00
2010s x Middle Selective	(1.61)	0.20	(8.06)	0.00
Dispersion x 2000s x Highly Selective	(306.61)	8.11	(37.79)	0.00
Dispersion x 2010s x Highly Selective	(529.64)	9.83	(53.86)	0.00
Dispersion x 2000s x Middle Selective	(70.81)	6.62	(10.69)	0.00
Dispersion x 2010s x Middle Selective	(135.59)	7.86	(17.25)	0.00

Table 4: Time-varying model coefficients

The R-squared of this regression indicates that about 0.8% of the variability in alpha can be explained by variability in the predictors. The model now explains about 3 times as much of the variance in alpha as before. Again, this low R-squared is likely due to the noisy nature of monthly alphas mentioned earlier. VIX remains a significant negative predictor of alpha; a 1-point increase in VIX is associated with an 18 basis point decrease in alpha. Due to interaction effects between the other relationships being included in the model, it makes sense to interpret each of those relationships by decade.

Table 5 displays coefficients for intercept and slopes associated with dispersion and VIX for each of the fund categories. Each column corresponds to funds in a particular decade and of a particular selectivity. The mean alpha row can be interpreted as average fund alpha at mean levels of VIX and dispersion, measured as a percentage. Mean dispersion level is calculated for each decade and indicated with a vertical gray line on each graph. The dispersion coefficient row corresponds to the model coefficient for dispersion for each fund type. It can be interpreted as the basis point change in alpha associated with a 1 percentage point increase in dispersion.

Decade		1990s			2000s			2010s	
Selectivity	Low	Middle	High	Low	Middle	High	Low	Middle	High
Mean Alpha	(0.01)	0.54	2.29	0.44	(0.04)	(0.01)	(0.86)	(0.90)	(1.14)
Dispersion Coefficient	(33.82)	40.95	261.53	50.55	54.52	39.29	(14.64)	(75.46)	(248.94)

Table 5: Mean alpha and dispersion coefficient for each fund category

4.2.1) 1990s

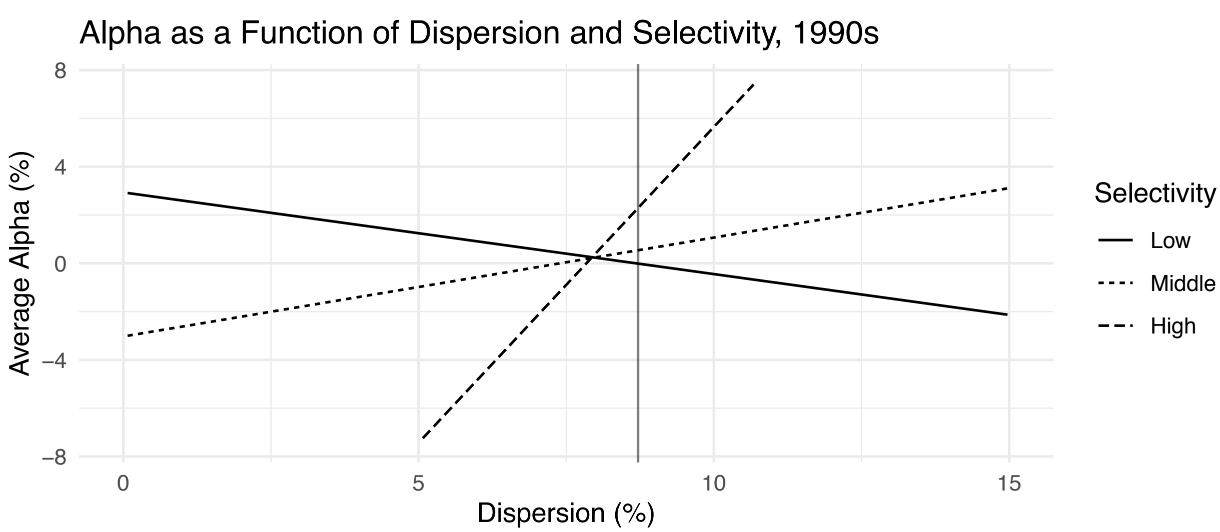


Figure 5: 1990s time-varying model visualization

In the 1990s, low selectivity funds averaged an alpha of -0.01% at mean levels of dispersion and VIX, compared to 0.54% and 2.29 % for middle and high selectivity funds respectively. These differences are statistically significant. Each of the selectivity groups had significant and large differences in the relationship between dispersion and alpha. Highly selective firms had the largest positive relationship between dispersion and alpha: a one percentage point increase in dispersion is associated with an increase of 260 basis points in monthly alpha. For middle and low selectivity funds, the expected change in alpha for a one percentage point increase in dispersion is positive 41 basis points and negative 34 basis points respectively. In other words, highly selective funds outperformed their peers in the 1990s, especially during times of high dispersion. This evidence is consistent with the existence of "stock pickers' markets" in the 1990s, as the most active funds were able to outperform relative to what was expected when there existed high variability in stock returns.

4.2.2) 2000s

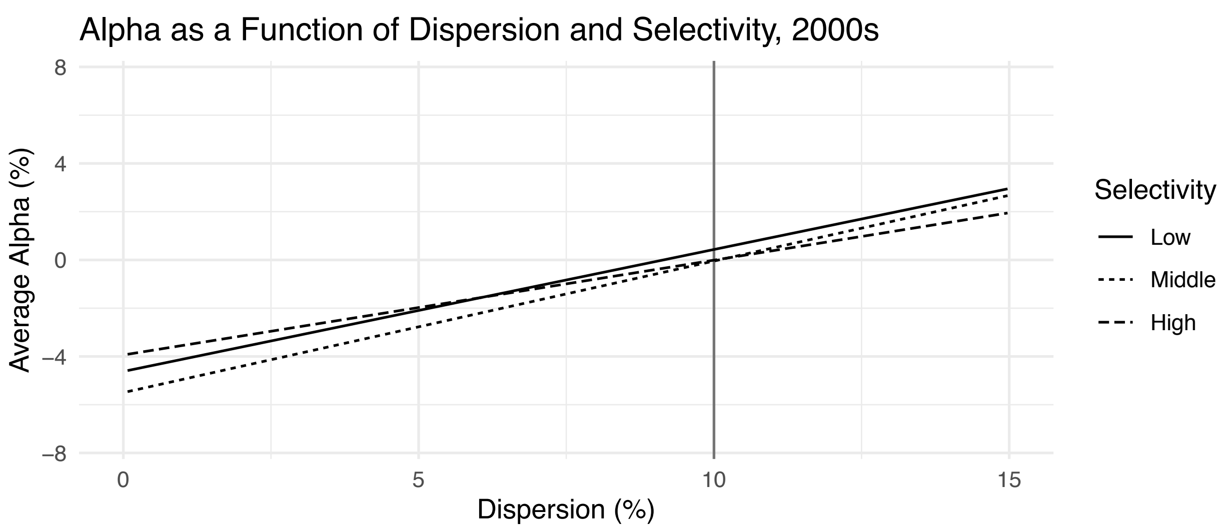


Figure 6: 2000s time-varying model visualization

In the 2000s, low selectivity funds averaged an alpha of 0.44% at mean levels of dispersion and VIX, compared to -0.04% and -0.01% for middle and high selectivity funds respectively. For highly selective funds, a 1 percentage point increase in dispersion is associated with a 39 basis point increase in alpha, compared to expected 51 and 54 basis point increases for middle and low selectivity funds respectively. Low selectivity funds outperformed middle and high selectivity funds at mean levels of dispersion and had a very similar relationship between outperformance and dispersion during this period. This result seems surprising, given that general intuition is that more active funds would outperform during periods of crisis, and this period included the worst financial crisis in the last 75 years.

4.2.3) 2010s

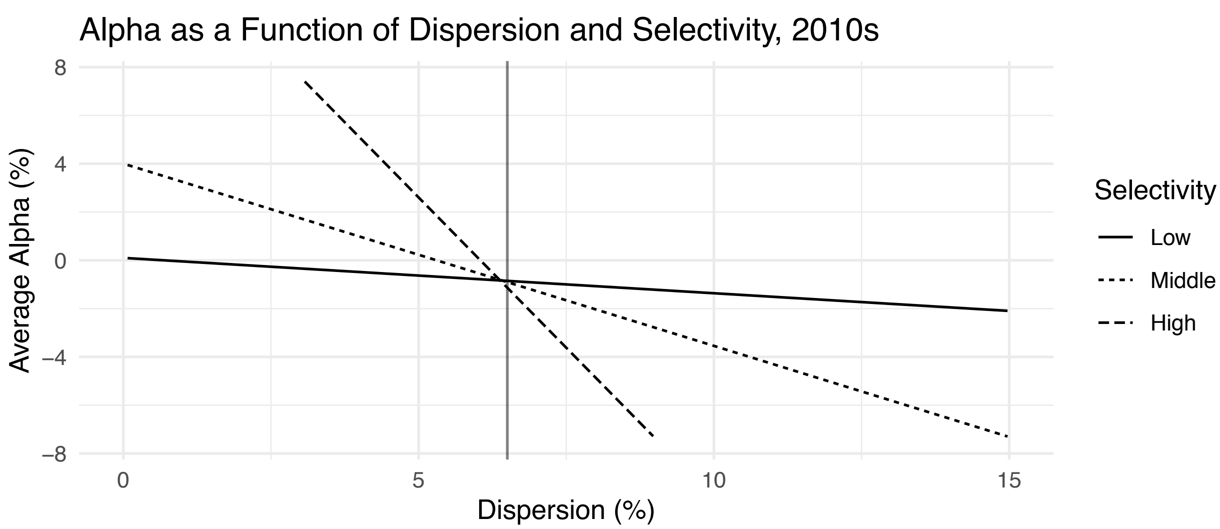


Figure 7: 2010s time-varying model visualization

In the 2010s, low selectivity funds averaged an alpha of -0.86% at mean levels of dispersion and VIX, compared to -0.90% and -1.14 % for middle and high selectivity funds respectively. These differences are not statistically significant. Each of the selectivity groups had large differences in the relationship between dispersion and alpha. For highly selective funds, a 1 percentage point increase in dispersion is associated with a 249 basis point decrease in alpha, compared to a 76 basis point decrease for middle selectivity funds and a 15 basis point decrease for low selectivity funds. These differences are statistically significant. In other words, highly selective funds underperformed relative to their peers during periods of high dispersion. This result may be driven by the low-volatility bull market during this period which favored passive vs. active investing.

4.3) Fund Dispersion vs Market Dispersion

I then tested whether dispersion in security returns is associated with dispersion in fund alpha. To do so, I estimated dispersion of fund alphas as cross-sectional standard deviation of fund alphas for every month in the data. I created two different models to explain fund dispersion with market dispersion: a simple linear model, and a model with log transformations of both fund dispersion and market dispersion. I chose to explore this log transformed model in addition to the linear model because the data displayed a fanning pattern, and thus violated the heteroskedasticity condition of linear regression. The two models are shown below, and their outputs are shown in tables 6 and 7. Scatterplots for each model are shown in figure 8.

Linear:

$FundDispersion_t = \beta_0 + \beta_1 Dispersion_t + \epsilon_t$

N	R-Squared	F	P-Value
312	0.118	41.29	0.000

	Estimate	Standard Error	T	P-Value
Intercept	0.01	0.00	6.73	0.00
Dispersion	0.10	0.02	6.43	0.00

Table 6: Linear dispersion vs. dispersion model output

Log-Log:

$log(FundDispersion_t) = \beta_0 + \beta_1 log(Dispersion_t) + \epsilon_t$

N	R-Squared	F	P-Value
312	0.295	129.8	0.000

	Estimate	Standard Error	T	P-Value
Intercept	(2.70)	0.13	(21.59)	0.00
Log(Dispersion)	0.56	0.05	11.39	0.00

Table 7: Log-log dispersion vs. dispersion model and output

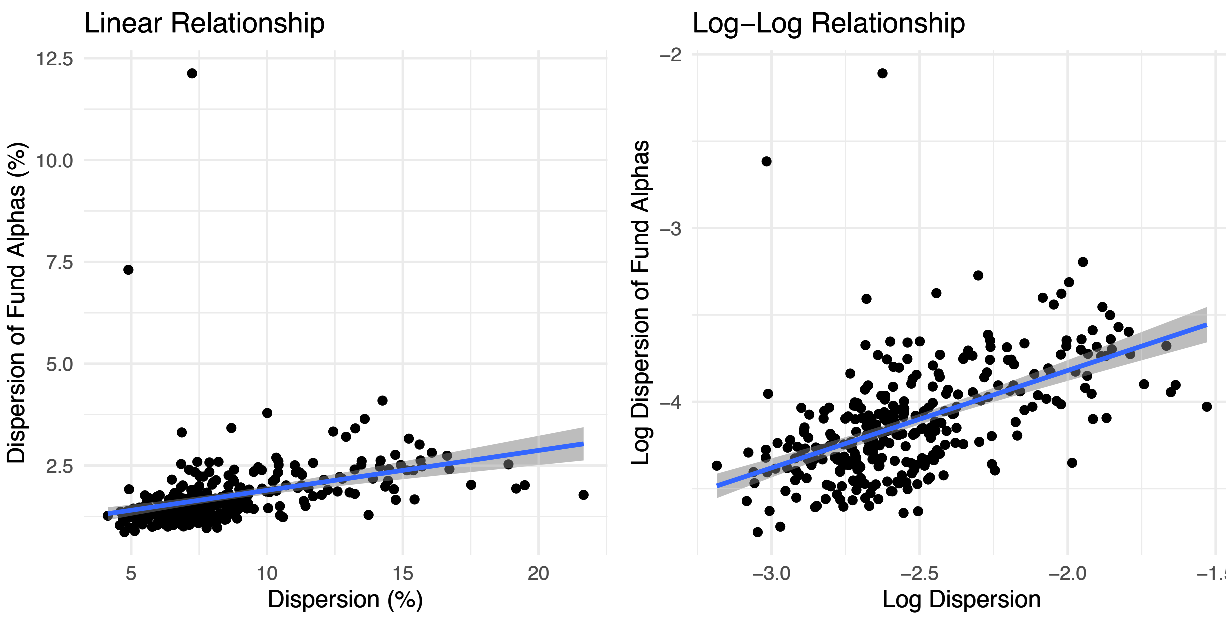


Figure 8: Scatterplots for linear and log-log models

Since the log-log model has an r-squared that is larger than that of the simple linear model (0.30 vs 0.12), and because it better satisfies the condition of heteroskedasticity (as seen in the figure above), I report all conclusions for the log-log model. This model indicates that market dispersion has a statistically significant positive relationship with dispersion of mutual fund alpha. A 10% increase in market dispersion is associated with about a 6% increase in dispersion of mutual fund alphas. This result is consistent with de Silva, Sapra and Thorley (2001), who demonstrated that market dispersion is associated with dispersion of mutual fund returns.

This result also suggests that it is unwise for an investor to make allocations to a particular fund based on that fund's alpha during a period of high market dispersion. Since times of high market dispersion are associated with high dispersion of fund alphas, it is difficult to discern whether that alpha is the result of manager skill or random chance.

5. Conclusions and Further Research

This paper suggests that dispersion can be a significant predictor of mutual fund alpha, but there are large differences in this relationship over time. During the 1990s, the most selective funds had a very significant positive relationship between dispersion and alpha, more so than their less selective peers. This result is intuitive and consistent with the idea of a stock picker's market; the most active stock pickers have the most to benefit from periods of high dispersion. In the 2000s, this relationship was positive for all funds, providing more evidence in favor of the existence of stock picker's markets. Additionally, high selectivity funds performed roughly in line with their peers at all levels of dispersion during this period. Since this period was dominated by two financial crises: the dot-com bubble and the global financial crisis of 2008, this result seems to contradict a popular selling point for active investing, that it is better to have money actively managed during times of turmoil. The 2010s provided the most surprising results, as funds of all selectivity levels demonstrated significant negative relationships between dispersion and alpha. This result is even more surprising when considering the explosive rise in the market share of passively managed investment vehicles. One might expect that a smaller pool of active managers would result in more mis-pricings in the market, allowing the remaining active managers to outperform, but this is not what the data have shown.

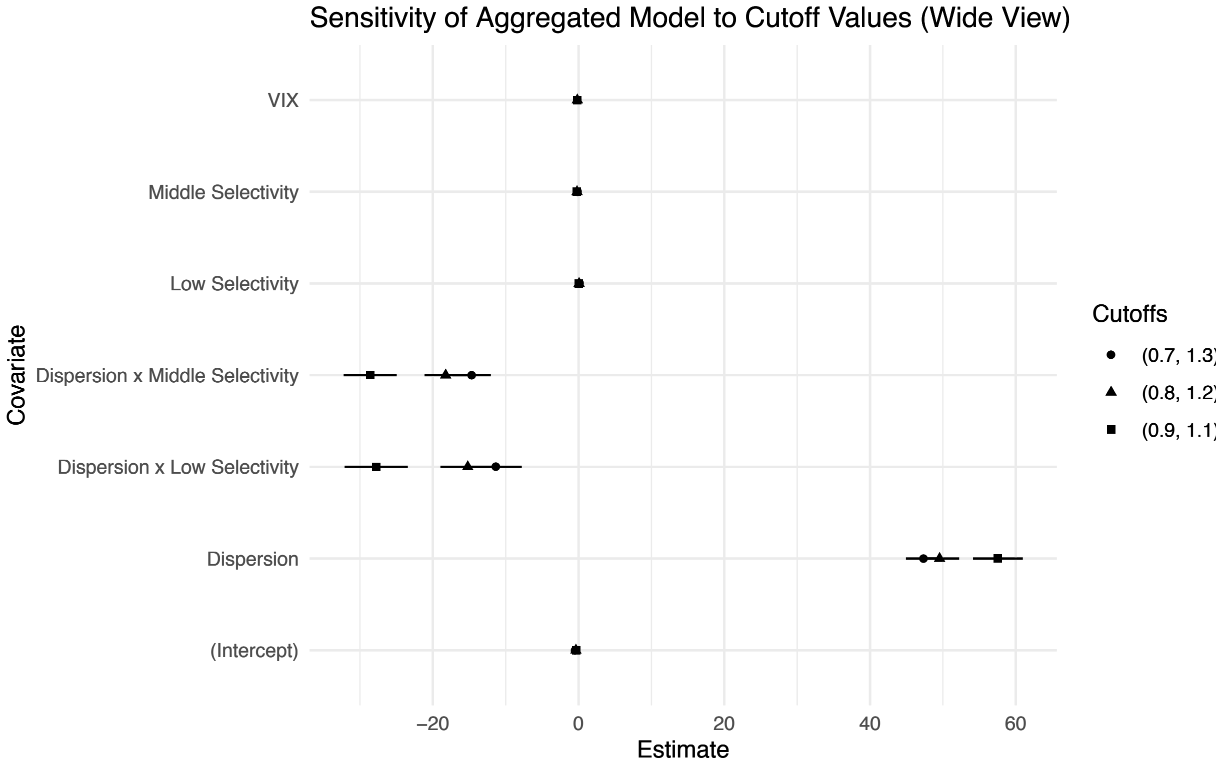
There are several possible explanations for these results. One possible explanation is that the stock market has simply gotten more efficient over time. In the 1990s, managers could have taken advantage of inefficient pricing in the market, and the most active managers would have the most opportunity to generate alpha. By the 2010s, there could have been fewer inefficiencies for active managers to exploit. Managers would incur transaction costs when picking stocks during times of high dispersion, but not generate outsized returns, resulting in a negative relationship between dispersion and alpha. This explanation relies on the assumption that funds would transact more often during times of high dispersion. This question could be interesting to explore in further research.

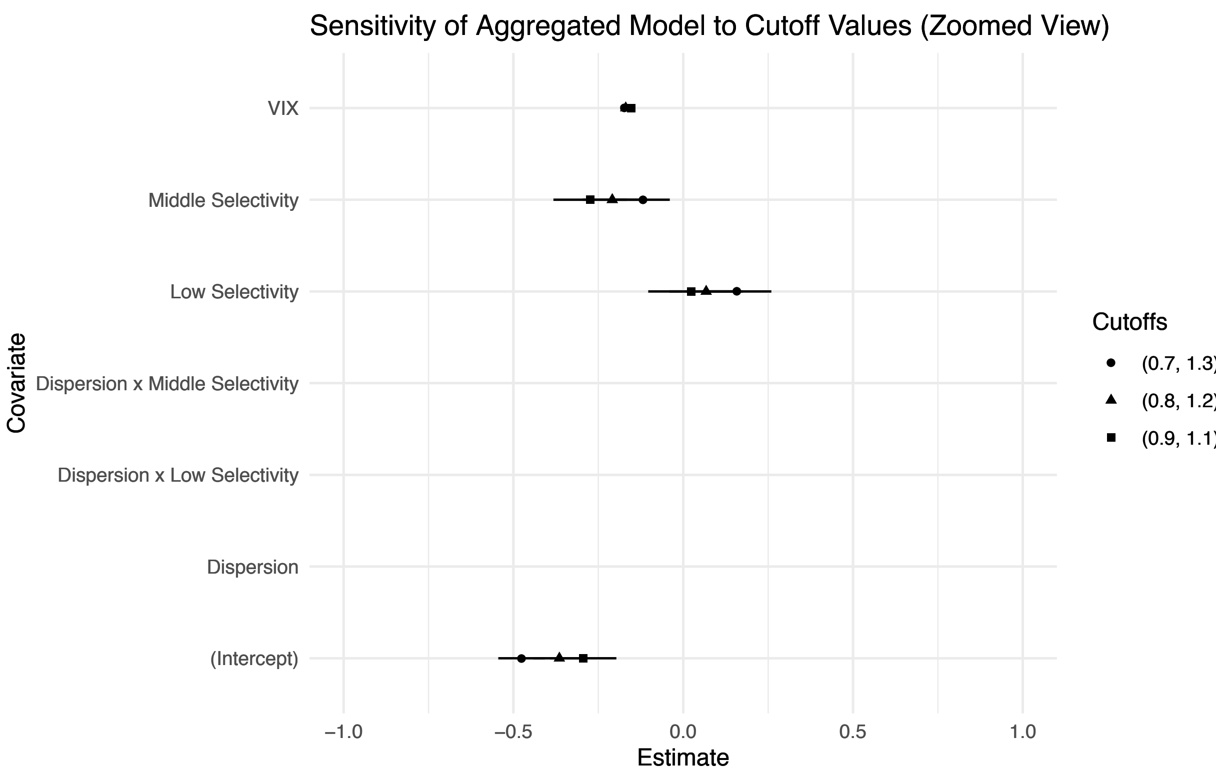
Another possible explanation for the results is that mutual fund managers have gotten less skilled relative to the average market participant over the past few decades. That is not to say that active managers have become less skilled overall, but the average market participant may have gotten more sophisticated. It could be the case that in the 1990s, active mutual funds had a stock picking advantage over less-sophisticated retail investors. Active funds in high dispersion markets would have the most opportunity to exploit this edge. As more sophisticated participants, such as quantitative hedge funds, have entered the market and grown in size in recent years[2], active managers could now be at a disadvantage relative to the average market participant. Funds which transact the most would therefore be expected to perform the worst, explaining the near complete reversal in trends between the 1990s and 2010s. Further evidence is needed to evaluate whether the average market participant has actually gotten more sophisticated, so this could be an interesting topic for future research.

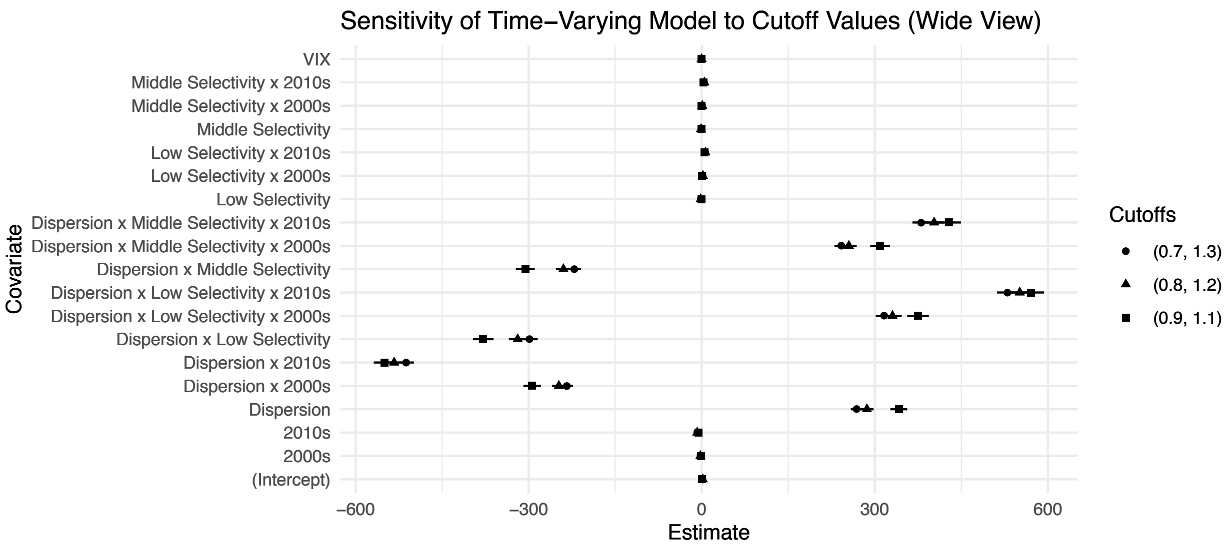
From a practical perspective, this analysis does not recommend using a fund's performance during times of high dispersion to determine whether to make allocations to a particular fund. While the results are statistically significant, they apply to funds on average, and are difficult to apply to a single fund. As seen by the particularly low R-squared values in the regressions, there is a huge amount of noise in monthly alphas. Furthermore, dispersion in the market is strongly correlated with dispersion of fund alphas, meaning it is very difficult to discern whether a particular fund's strong performance in a time of high dispersion is the product of managerial skill or luck.

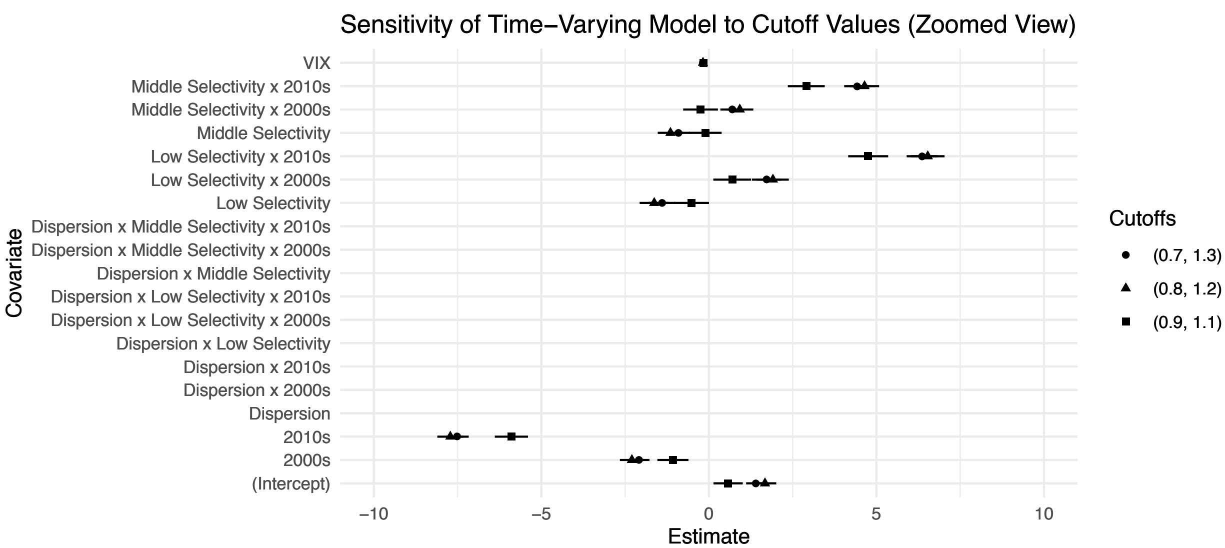
[2] https://www.ft.com/content/ff7528bc-ec16-11e7-8713-513b1d7ca85a

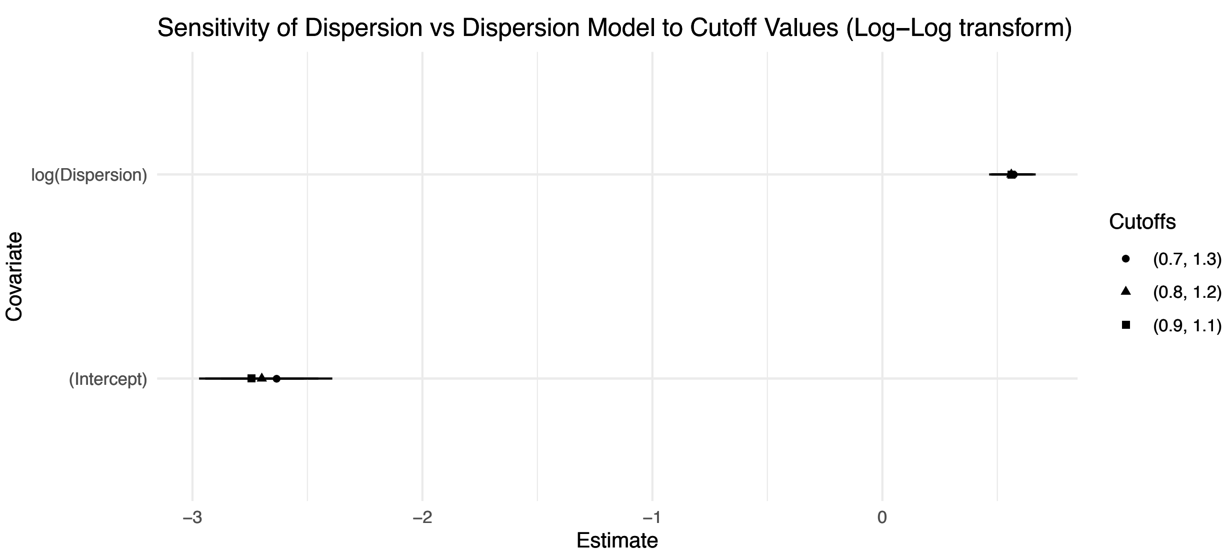
Appendix 1: Sensitivity of Model to Beta Cutoffs






Sensitivity of Time−Varying Model to Cutoff Values (Wide View)
VIX
Middle Selectivity x 2010s
Middle Selectivity x 2000s
Middle Selectivity
Low Selectivity x 2010s
Low Selectivity x 2000s
Low Selectivity
Dispersion x Middle Selectivity x 2010s
Dispersion x Middle Selectivity x 2000s
Dispersion x Middle Selectivity
Dispersion x Low Selectivity x 2010s
Dispersion x Low Selectivity x 2000s
Dispersion x Low Selectivity
Dispersion x 2010s
Dispersion x 2000s
Dispersion
2010s
2000s
(Intercept)
Covariate
−600
−300
0
300
600
Estimate
Cutoffs
(0.7, 1.3)
(0.8, 1.2)
(0.9, 1.1)


Sensitivity of Time−Varying Model to Cutoff Values (Zoomed View)
VIX
Middle Selectivity x 2010s
Middle Selectivity x 2000s
Middle Selectivity
Low Selectivity x 2010s
Low Selectivity x 2000s
Low Selectivity
Dispersion x Middle Selectivity x 2010s
Dispersion x Middle Selectivity x 2000s
Dispersion x Middle Selectivity
Dispersion x Low Selectivity x 2010s
Dispersion x Low Selectivity x 2000s
Dispersion x Low Selectivity
Dispersion x 2010s
Dispersion x 2000s
Dispersion
2010s
2000s
(Intercept)
Covariate
−10
−5
0
5
10
Estimate
Cutoffs
(0.7, 1.3)
(0.8, 1.2)
(0.9, 1.1)


Sensitivity of Dispersion vs Dispersion Model to Cutoff Values (Log−Log transform)
log(Dispersion)
(Intercept)
Covariate
−3
−2
−1
0
Estimate
Cutoffs
(0.7, 1.3)
(0.8, 1.2)
(0.9, 1.1)

References

Amihud, Y., & Goyenko, R. (2013). Mutual FundsR^2 as Predictor of Performance. *Review of Financial Studies*, *26*(3), 667–694. doi: 10.1093/rfs/hhs182

Ankrim, E. M., & Ding, Z. (2002). Cross-Sectional Volatility and Return Dispersion. *Financial Analysts Journal*, *58*(5), 67–73. doi: 10.2469/faj.v58.n5.2469

Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. *The Journal of Finance*, *52*(1), 57–82. doi: 10.1111/j.1540-6261.1997.tb03808.x

Carhart, M. M., Carpenter, J. N., Lynch, A. W., & Musto, D. K. (2002). Mutual Fund Survivorship. *Review of Financial Studies*, *15*(5), 1439–1463. doi: 10.1093/rfs/15.5.1439

Cremers, K. J. M., & Petajisto, A. (2009). How Active Is Your Fund Manager? A New Measure That Predicts Performance. *Review of Financial Studies*, *22*(9), 3329–3365. doi: 10.1093/rfs/hhp057

Davis, J. L. (2001). Mutual Fund Performance and Manager Style. *Financial Analysts Journal*, *57*(1), 19–27. doi: 10.2469/faj.v57.n1.2416

de Silva, H., Sapra, S., & Thorley, S. (2001). Return Dispersion and Active Management. *Financial Analysts Journal*, *57*(5), 29–42. doi: 10.2469/faj.v57.n5.2479

Fama, E. F., & French, K. R. (1992). The Cross-Section of Expected Stock Returns. *The Journal of Finance*, *47*(2), 427. doi: 10.2307/2329112

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3–56. doi: 10.1016/0304-405x(93)90023-5

Frazzini, A., Friedman, J., & Pomorski, L. (2016). Deactivating Active Share. *Financial Analysts Journal*, *72*(2), 14–21. doi: 10.2469/faj.v72.n2.2

Jensen, M. C. (1968). The Performance of Mutual Funds in the Period 1945-1964. *The Journal of Finance*, *23*(2), 389. doi: 10.2307/2325404

Jones, R. C., & Wermers, R. (2011). Active Management in Mostly Efficient Markets. *Financial Analysts Journal*, *67*(6), 29–45. doi: 10.2469/faj.v67.n6.5

Low, C. (2000). The Fear and Exuberance from Implied Volatility of S&P 100 Index Options. *SSRN Electronic Journal*. doi: 10.2139/ssrn.194288

Petajisto, A. (2013). Active Share and Mutual Fund Performance. *Financial Analysts Journal*, *69*(4), 73–93. doi: 10.2469/faj.v69.n4.7