JOIM
www.joim.com

Journal Of Investment Management, Vol. 15, No. 1, (2017), pp. 50–74

PICKING "WINNER" FUNDS*

Joshua Livnat[a,b], *Gavin Smith*[b] *and Martin Tarlie*[b]

One of the most crucial decisions for investors and plan sponsors is the selection of funds among the thousands of available alternatives. We find that regardless of the initial criterion used to rank funds based on past performance, more diversified top funds outperform concentrated top funds in the subsequent year. This better performance is attributed to the more consistent returns of funds with diversified holdings. We also find that when the initial criterion is based on a manager's skill, as measured by a positive intercept in a regression of past fund returns on the five Fama–French factors, active share is a useful tool to predict future winner funds among top skill managers. However, diversified top funds provide slightly higher returns and less severe drawdowns than funds with high active share and top manager skill.

To account for the problem that high information ratio can be associated with low return but even lower tracking error, we introduce the Modified Information Ratio (IR) measure. This measure adjusts the conventional IR to account for an investor's desired alpha. The Modified IR measure and conventional IR behave similarly with respect to diversification. We find that top Modified IR funds that are also diversified—winner funds—have significantly better future 12-month returns than top Modified IR funds that are concentrated.

Many investors and financial intermediaries grapple with the selection of funds among the thousands available for investment. Investors and their financial advisors attempt to construct a sensible portfolio of investments from the bewildering array of available options. Pension plan administrators identify funds for their own portfolios or for their participants. Similarly, consultants and advisors select a short list of recommended fund strategies for their clients. A crucial element in these decisions is the prediction of a fund's future performance, a topic that has attracted a long line of academic and practitioner research.

*These materials represent the views and opinions of the authors regarding the economic conditions, asset classes, or financial instruments referenced herein and are not necessarily the views of QMa.

[a]Professor Emeritus of Accounting, Stern School of Business Administration, New York University, USA. E-mail: jlivnat@stern.nyu.edu

[b]Quantitative Management Associates, 2 Gateway Center, Newark, NJ 07102, USA.

A variety of measures have been shown to be effective in selecting funds that subsequently outperform their benchmark. These measures include past performance (i.e. momentum), tracking error, information ratio, various Fama–French regression statistics, and active share. We refer to the best funds in each of these categories as top funds. Within each category of top funds, there is a great deal of variation in fund characteristics. Our focus in this paper is on the difference between those top funds that are diversified and those that are concentrated.

Our evidence suggests that regardless of which measure is used to rank funds, top funds that are more diversified (i.e. funds that have more portfolio holdings) outperform their less diversified top fund peers. In fact, unlike prior evidence that the average fund net return in excess of the benchmark is negative, we show that for some measures the most diversified funds actually have a **positive** average net return above their benchmark in the subsequent year. Furthermore, these funds have a 50% chance of outperforming their benchmark in the subsequent year. In contrast, the concentrated top fund performers have **negative** average net returns above their benchmark in the subsequent year, and only 43% of them are likely to outperform their benchmark in the subsequent year. Our preferred measure is the information ratio (IR). This measure is attractive because it combines past performance (momentum) and consistency (tracking error) into a single measure.

However, the shortcoming of IR is that it may tilt us to funds that "hug" their benchmark very closely, delivering a tiny net return in excess of the benchmark with a commensurate tiny tracking error. Such funds, although superior from an IR perspective, may not enable the investor to have sufficient return to pay for consumption. Borrowing the colloquialism "you can't pay your mortgage with Sharpe Ratio" to the case of IR, summarizes this notion. At the extreme, a passive investment in the benchmark offers a negative IR due to the management fees associated with this investment.

We address this issue of low returns but even lower tracking error by modifying the IR to account for an investor's desired alpha. The new measure—Modified IR—is defined as the difference between the fund's alpha (net return above the benchmark) and the investor's desired alpha, divided by the fund's tracking error. The attractive feature of this measure is that, *ceteris paribus*, a higher desired alpha favors higher tracking error. It is important to recognize that the investor's desired alpha represents an investor preference; it is not a parameter to be optimized.

Empirically, we find that top funds based on Modified IR have tracking error, turnover and expense ratio that are positively correlated with the desired alpha. Furthermore, consistent with the other measures, we find that diversified top Modified IR funds outperform concentrated top Modified IR funds.

Most of the measures used to define top funds are uncorrelated with the degree of diversification. For these measures, "winner funds" are the diversified funds within the select group of the measure's top funds. However, the top funds for active share and stock picking skill (i.e. the *t*-statistic of the Fama–French regression intercept) are inherently concentrated. It therefore does not make sense to compare diversified and concentrated funds for these two measures. For these inherently concentrated measures, we find that funds are best selected by first picking the top funds based on the *t*-statistic of the Fama–French regression intercept, and then within that group selecting the funds with high active share.

Thus, our study identifies two different approaches to select funds that are likely to outperform in the future: (i) diversified funds within the top funds

defined by measures uncorrelated with diversification, and (ii) high active share funds within the top funds defined by stock picking skill. The main difference between these two approaches is that the first approach leads to more diversified funds, typically exposed to risk factors such as value, quality, or size, and the second approach leads to more concentrated funds with higher idiosyncratic volatility. While our results show that the approach based on diversification has had slightly higher returns, slightly greater likelihood of beating their benchmark in the subsequent year, and a lower drawdown, these differences are not statistically significant.

Our study contributes to the literature along several dimensions. We present two different paths to picking winner funds, a diversified approach and a concentrated one. Furthermore, we suggest a new measure—Modified IR—to rank funds that combines past performance, tracking error, and an investor's desired alpha. Both of these contributions could be used in future academic works, as well as by professionals in the investment community.

Sections 1–3 compare the future performance of funds selected according to the various measures in the literature,[1] and examines the performance of the additional criteria of active share and fund diversification. Section 4 introduces the Modified IR measure and describes how it is estimated. It also provides evidence about the superiority of more diversified funds relative to more concentrated ones in the context of the Modified IR method. The last section summarizes the findings and concludes the study.

1 Data and results

1.1 Data and variables

This study is based on the CRSP Survivor-Bias-Free US Mutual Fund database, accessed through WRDS. As in prior studies, we focus on domestic equity funds which are neither sector funds, nor index funds, and which have a minimum net asset value of $15 million.[2] For each of these funds we identify all the funds that belong to the same CRSP Class Group, and use weighted average variables for our analyses, where the class weights are based on the ratio of the latest Total Net Assets (TNA) in a class to the sum of TNA over all classes of funds within the same CRSP class group. At the end of each December, we use the prior 36 monthly weighted average returns to calculate several measures that are used to select funds for the following calendar year.[3] The CRSP database provides net returns for the funds, which is what we use in this study. We eliminate funds that do not have the full 36-month history of returns. We examine fund cumulative returns over the subsequent year. If a fund stopped trading during the year, we use the cumulative returns through the month that it stopped trading.

One of the difficulties of working with the CRSP mutual fund database is the selection of benchmark for each fund. We use the benchmark selected by Petajisto (2013) in his study, which he primarily obtained from the fund prospectus (p. 78).[4] We use that same fund benchmark in earlier or subsequent years not covered by Petajisto's dataset. For funds that are not covered by Petajisto's dataset, we use an active share dataset created by Nomura that was prepared using the CRSP mutual fund database and the fund holdings reported by CRSP.[5] There is a short period where both datasets overlap. We examined the consistency of the active share estimates for these observations. We found that the two datasets selected the same benchmark only in about 50% of the cases, indicating how difficult it is to find the proper benchmark for a fund. However, the correlation between the active share estimates in the two datasets is about 75%. For funds that

did not have benchmark information in either of these datasets we used the S&P 500 Index as the benchmark. All benchmark returns include dividends.

To predict future fund performance, we use many of the measures that are suggested in the literature. All of these measures (except active share, which we did not compute ourselves) are based on data taken from the previous 36 months. The first measure is momentum, defined as the fund's cumulative 36-month return in excess of the benchmark. The second measure is the tracking error, defined as the standard error of the differences between the fund and the benchmark log monthly returns over the prior 36 months. The third measure is information ratio (IR), defined as the mean of the differences in the monthly log returns between the fund return and the benchmark return over the prior 36 months, scaled by the tracking error. For the measures based on the Fama–French statistics, we use a five-factor model (market return in excess of the risk-free rate, high minus low *B/M*, small minus large cap, 11-month momentum and short-term reversal).[6] For each fund, we regress the prior 36 monthly returns of the fund, in excess of the risk-free rate, on the Fama–French five-factor returns. The fourth measure is the intercept of this regression and measures the unique contribution of the fund manager beyond the known risk factors. The fifth measure is the *t*-statistic of this intercept and is a measure of the significance of the ability to generate returns that are not explainable by the five Fama–French factors. The sixth measure is the *R*-Square of the regression of the fund return on the five Fama–French factors. As pointed by Amihud and Goyenko (2013) this measure captures the selectivity of the fund manager, and can be used in predicting future performance. Note that all the above measures are return-based and do not require any knowledge of individual fund holdings. Finally, the seventh measure is active share, which we obtain from Petajisto and Nomura. Among the fund characteristics we examine are the weighted average expense ratio and the fund turnover as reported in the CRSP database. We also use the number of holdings in a fund to distinguish between funds with more diversified and more concentrated approaches. These data are from the CRSP mutual fund data, and are available from 2002 onwards.

2 Summary statistics

Table 1 reports summary statistics for all the fund-year observations in the study. As Figure 1 shows, the data spans the years 1961–2013, with just one fund in 1961, three in 1962, over 100 in 1969 and over 1,000 since 1994. As Panel A in Table 1 shows, the average fund has a mean annual, net of fees, return in the year after fund selection of −131 BP, consistent with prior findings that the average fund detracts from investors' value over time. The average fund had an expense ratio of 120 BP, which means that most of the fund underperformance is due to fund expenses.[7] Funds also vary in their turnover ratios from 15% for the 10th percentile to 171% for the 90th percentile. We also see considerable variation in the number of holdings with a mean of 128, but a median of only 78 positions.

Turning to Panels B and C of Table 1, we see that funds in our sample, which requires that the prior 36 monthly returns be available, have higher average returns in the subsequent year: −116 BP compared to −150 BP for those that are not included in our sample. Furthermore, we see that the sample funds have lower expenses and turnover ratios than non-sample funds. These lower expenses and turnover contribute to their better net performance. Finally, we also observe that sample funds are larger in terms of net assets under management but have slightly fewer positions in their portfolios. They are similar to

Table 1 Summary statistics.

Variable	*N*	Mean	Median	10%	90%	St. Dev.
Panel A: All funds						
Annual Excess Return (BP)	50,401	(131)	(137)	(1,120)	840	1,019
Three-Year Excess Return (BP)	35,914	(437)	(409)	(2,774)	1,740	2,196
Expense Ratio (BP)	44,821	120	116	63	185	51
Turnover Ratio (%)	41,712	88	63	15	171	126
Assets ($ million)	50,401	752	129	18	1,361	3,199
Active Share (%)	31,666	77	82	54	97	19
Number of Holdings	21,055	128	78	31	271	160
Panel B: Not in our sample						
Annual Excess Return (BP)	21,861	(150)	(146)	(1,238)	882	1,102
Three-Year Excess Return (BP)	15,191	(531)	(481)	(3,042)	1,732	2,304
Expense Ratio (BP)	18,056	124	121	50	200	58
Turnover Ratio (%)	15,579	95	65	14	180	152
Assets ($ million)	21,861	283	66	14	592	989
Active Share (%)	10,168	77	82	52	97	21
Number of Holdings	6,934	136	77	27	330	179

Source: CRSP Survivor-Bias-Free US Mutual Fund database, December 2014, http://www.petajisto.net/data.html, Nomura Securities, http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library. html. *Note*: This table reports summary statistics for the domestic equity, non-index, and non-sector fund-observations in the CRSP mutual funds database with net assets above $15 million (Panel A), for the funds that were not included in our sample because of lacking 36 months of returns (Panel B), and for the funds that were included in our testing sample (Panel C). Funds are included in the sample each December, and their performance is measured in the subsequent year(s). *Annual Excess Return* (*BP*): The fund's annualized subsequent 12-month return over the months January–December, measured in basis points in excess of a benchmark, i.e. the cumulative return, net of fees, on the fund in the next 12 months minus the cumulative return on the benchmark (e.g., the S&P 500 Index) in the same period. *Three-Year Excess Return* (*BP*): The fund's subsequent 36-month return measured in basis points in excess of the benchmark, i.e. the cumulative return on the fund in the subsequent 36 months minus the cumulative return on the benchmark in the same period. *Expense Ratio* (*BP*): Funds' expense ratio measured in basis points as reported by the CRSP mutual fund data. *Turnover Ratio* (%): Funds' turnover ratio obtained from the CRSP mutual fund data. *Assets* (*$ Million*): Funds' total net assets obtained from the CRSP mutual fund data. *Active share* (%): Active share of a fund relative to its benchmark index as of each report date; obtained from Petajisto or Nomura. *Number of Holdings*: The number of holdings in a fund; obtained from the CRSP mutual fund data. *36-Month Momentum*: Cumulative return of the fund over the past 36 months minus the cumulative return on the benchmark for the same time period. *36-Month TE* (*BP*): The tracking error measured in basis points is the standard error of the differences between the fund and the benchmark log monthly returns during the prior 36 months. *Information Ratio* (*BP*): Information ratio measured in basis points is the mean of the differences in the monthly log returns between the fund return and the benchmark return over the prior 36 months, scaled by the tracking error. *36-Month FF Regression Intercept, t-Statistics, R-Square*: For each fund, we regress the monthly returns of the fund, in excess of the risk-free rate, on the FF five-factor returns over the prior 36 months. The intercept of this regression measures the unique contribution of the fund manager beyond the five known risk factors. The *t*-statistic of this intercept is a measure of significant ability to generate returns that are not explainable by the traditional FF factors. Finally, we use the *R*-Square of the regression of fund returns in the prior 36 months on the five FF factors as a measure of manager's skill.

Table 1 (*Continued*)

Variable	*N*	Mean	Median	10%	90%	St. Dev.
Panel C: Sample funds						
Annual Excess Return (BP)	28,540	(116)	(129)	(1,052)	808	950
Three-Year Excess Return (BP)	20,723	(367)	(356)	(2,577)	1,744	2,111
Expense Ratio (BP)	26,765	117	114	68	172	45
Turnover Ratio (%)	26,133	83	62	16	166	107
Assets ($ Million)	28,540	1,111	214	31	2,083	4,126
Active Share (%)	21,498	78	82	55	96	19
Number of Holdings	14,121	124	78	34	248	149
36-Month Momentum	28,540	(317)	(359)	(2,656)	1,908	2,555
36-Month TE (BP)	28,540	194	160	77	351	159
Information Ratio (BP)	28,540	(654)	(592)	(3,456)	2,059	2,214
36-Month Regression:						
Intercept	28,540	(0.069)	(0.059)	(0.476)	0.341	0.397
Intercept *t*-Statistic	28,540	(0.359)	(0.348)	(1.949)	1.223	1.270
R-Square	28,540	0.914	0.936	0.831	0.981	0.085
Panel D: Sample funds with rank and active share (2002–2013)						
Annual Excess Return (BP)	12,023	(82)	(102)	(738)	580	636
Three-Year Excess Return (BP)	7,823	(310)	(313)	(1,717)	1,089	1,283
Expense Ratio (BP)	10,873	118	115	75	164	41
Turnover Ratio (%)	10,787	81	62	18	161	75
Assets ($ Million)	12,023	1,263	238	32	2,191	4,808
Active Share (%)	12,023	77	80	56	95	164
Number of Holdings	12,023	123	79	36	242	144
36-Month Momentum	12,023	(257)	(284)	(1,717)	1,235	1,361
36-Month TE (BP)	12,023	151	131	69	251	93
Information Ratio (BP)	12,023	(635)	(565)	(3,211)	1,826	2,007
36-Month FF Regression:						
Intercept	12,023	(0.097)	(0.087)	(0.427)	0.239	0.301
Intercept *t*-Statistic	12,023	(0.473)	(0.449)	(2.025)	1.011	1.209
R-Square	12,023	0.936	0.951	0.871	0.984	0.056

non-sample funds in terms of active share. Examining the prior 36-month mean cumulative return of sample funds in Panel C, we can see that it is negative at −317 BP over the three-year period. The associated tracking error over the past three years is 194 BP, and the Information Ratio is also negative on average at −6.54%. Similarly, the five-factor Fama–French regression intercept in the prior 36 months is negative −0.069, indicating that the average fund lost about 69 BP after controlling for the five Fama–French factors. The intercept's *t*-statistic is also negative on average, but insignificantly different from zero for most funds. Finally, there is not much variation in the *R*-Square of the five-factor Fama–French regressions over the prior three years.

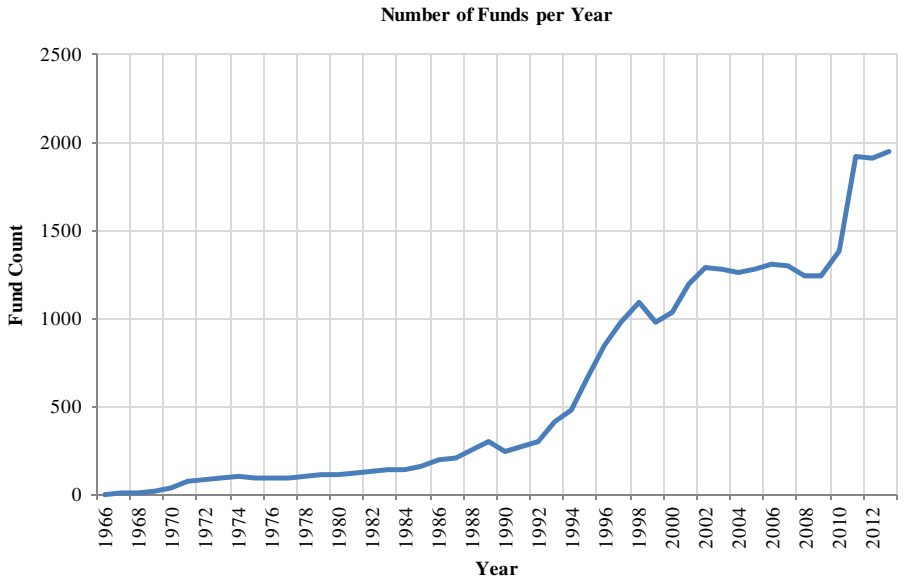

Figure 1 *Source:* CRSP Survivor-Bias-Free US Mutual Fund database; December 2014.

In what follows, we restrict the sample funds even further to those where we have not only the history of 36 monthly returns, but also have both active share and number of positions held in the portfolio. Because data about the number of positions is available in CRSP only from 2002 onwards, the final sample has 12,203 fund-year observations. The summary statistics for this restricted sample are very similar to those reported in Panel C of Table 1, and are reported in Panel D. There are some minor differences in the 36-month momentum, tracking error, IR, and the five-factor Fama–French regression intercept, but those differences are probably due to market conditions prevailing during the 2002–2013 period and not due to the data restrictions about availability of active share number of portfolio holdings. Due to the evolution of the mutual fund industry and data availability, we restrict our analysis and reporting of results in the body of the paper to the sample of funds in Section D of Table 1, i.e. those funds that had both number of positions reported in CRSP and active share (as well as prior 36 monthly returns) for the years 2002–2013. In a section about robustness checks below we describe the results for the larger sample which also spanned more years.

3 Top-quartile performance for various measures

Table 2 provides information about the returns (above the benchmark) of funds in the year after fund selection, segregated between the top quartile and all other funds. For most measures, the best quartile is the quartile with the highest scores for that year, except for the tracking error and the Fama–French regression *R*-Square, where the best quartile is the quartile with the lowest scores; lowest tracking error and the lowest *R*-Square, indicating a manager's skill in obtaining returns that are uncorrelated with the five Fama–French factors.

As can be seen in the table, the mean and median excess returns, i.e. return above the benchmark in the subsequent year, for the top funds (i.e. best quartile) are better than the other funds for all measures, except for the Fama–French regression *R*-Square, which has the opposite relationship.

Table 2 Fund net return in excess of its benchmark (BP) 2002–2013.

Measure		*N*	Mean (BP)	Median (BP)	Beat (%)	% Beat by 1%	% Miss by 1%	10th % (BP)
Momentum	Best quartile	3,031	−69	−71	45	37	48	−796
	Other	8,992	−86	−109	39	30	51	−720
Tracking Error	Best quartile	3,118	−81	−80	38	26	48	−499
	Other	8,905	−82	−111	42	34	51	−821
Information Ratio	Best quartile	2,983	−59	−62	45	36	47	−747
	Other	9,040	−89	−111	39	30	51	−736
FF Regression Intercept	Best quartile	3,070	−50	−83	43	35	49	−758
	Other	8,953	−92	−106	40	31	51	−731
FF Intercept *t*-Statistic	Best quartile	3,055	−56	−86	42	35	49	−747
	Other	8,968	−90	−106	40	31	51	−736
FF Regression *R*-Square	Best quartile	2,907	−122	−155	40	33	54	−1040
	Other	9,116	−69	−89	41	31	49	−656
Active Share	Best quartile	2,984	−60	−93	44	37	49	−960
	Other	9,039	−89	−105	40	30	50	−680
Number of Positions	Best quartile	3,018	−37	−39	44	32	44	−590
	Other	9,005	−96	−127	40	32	52	−780
All		12,023	−82	−102	41	32	50	−739

Source: CRSP Survivor-Bias-Free US Mutual Fund database, December 2014, http://www.petajisto.net/data.html, Nomura Securities, http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

Note: This table shows information about the returns (above the benchmark) of funds in the year after fund selection, segregated between the top quartile and all other funds for several measures. It includes data for funds that have returns for the prior 36 months, active share, and number of positions on CRSP for 2002–2013. *Momentum*: Cumulative return over the past 36 months minus the cumulative return of the benchmark over the same period. *Tracking Error*: The tracking error is the standard error of the differences between the fund and the benchmark log monthly returns over the prior 36 months. *Information Ratio*: Information ratio is defined as the mean of the differences in the monthly log returns between the fund return and the benchmark return over the prior 36 months, scaled by the tracking error. *Intercept FF, t-Statistics FF, and R-Square FF*: For each fund, we regress the monthly returns of the fund, in excess of the risk-free rate, on the FF five-factor returns over the prior 36 months. The intercept of this regression measures the unique contribution of the fund manager beyond known risk factors. The *t*-statistic of this intercept is a measure of significant ability to generate returns that are not explainable by the traditional FF factors. Finally, we use the *R*-Square of the above regression as a measure of the manager's skill. *Active share* (%): Active share of a fund relative to its benchmark index as of each report date; obtained from Petajisto or Nomura.*Number of Positions*: The number of holdings in a fund; obtained from the CRSP mutual fund data. Beat is the percentage of funds that are able to have returns in excess of their benchmark for the following year. % Beat (Miss) by 1% is the percentage of observations that beat (miss) their benchmark by more than 1% in the subsequent year. 10th % shows the excess return during the following year for the fund that ranked in the 10th percentile of all funds in that category.

Similarly, the percentage of funds that beat their benchmark in the subsequent year is higher for the top funds than for the other funds, again except for tracking error and the Fama-French Regression *R*-Square. We also examine a potential drawdown by looking at the 10th percentile next-year excess return. We typically find that this measure of drawdown is less favorable for funds in the best quartile as compared to all other funds, except for tracking error and the number

of positions held in the fund. This seems intuitive because the latter two measures are designed to reduce tracking error and offer greater diversification, both of which reduce the chance of large future losses. The table also shows the percentage of funds that beat or miss the benchmark in the subsequent year by more than 1%. As expected, funds that take bigger bets (high active share, for example) are more likely to beat the benchmark by more than 1%. Similarly, those that are more diversified (as measured by number of positions) are less likely to underperform by more than 1%.

A further examination of Table 2 shows some other interesting findings. The most noticeable is the superiority of the top funds that have the largest number of holdings; the "Number of Positions" line in Table 2. These funds, which are also the most diversified, have the highest next-year return. This group also has the second smallest drawdown among all the measures, indicating the performance consistency that can be expected from having a diversified portfolio. Next, contrary to prior studies in the literature, funds in the top quartile of active share deliver returns that are similar in magnitude to those delivered by IR, momentum and the Fama–French regression intercept. However, funds with the highest active share had a much larger drawdown than most other measures, consistent with the intuition that a high active share implies greater active bets, exposing the portfolio to greater future drawdown risk.

These results raise the question of whether a combination of measures can enhance the prediction of future fund performance, and if so, which measures should be combined. Intuitively, it makes sense to combine momentum and tracking error into IR, and possibly add another layer of safeguard to enhance consistency of returns through diversification. Similarly, one can enhance the identification of a manager's skill through the five-factor Fama–French regression intercept *t*-statistic with active share to identify managers who are both skillful and take bolder bets.

The results in Table 3 address this question of combining top funds and either diversification or active share. As in Table 2, there are 3,031 fund-year observations in the top quartile of momentum with an average net return of −69 BP. First, let's focus on the combination of momentum and active share. When we intersect the top momentum funds with those funds that are also in the top and bottom quartiles of active share we see average returns of −45 BP for the high active share funds, compared to a **positive** 24 BP for the lowest active share funds (although the difference between the two groups is not statistically significant).[8] Similar to what we saw in Table 2, the high active share funds had a larger drawdown, with −900 BP for the fund at the 10th percentile, as compared to −590 for the low active share and top momentum quartile.

A more interesting pattern emerges when we examine the performance of the intersection of top momentum funds with diversified (top quartile of holdings) and concentrated funds (bottom quartile of holdings).[9] In this case, we see that diversified top momentum funds have a net **positive** average return of 7 BP, as compared to a negative return of −88 BP for the concentrated funds (with the differences statistically significant). Furthermore, 50% of diversified top momentum funds beat their benchmark in the subsequent year, compared to 42% for the concentrated funds (with the differences statistically significant). Similarly, 51% of the concentrated funds miss their benchmark by more than 1% over the subsequent year, compared to only 40% for the diversified funds. Also, diversified funds at the 10% level of all funds had a negative return of −580 BP, as compared to −980 BP for the concentrated funds and −800 BP of all funds in the top momentum

Table 3 Future excess returns for groups based on prior 36-month measures, active share and number of positions 2002–2013.

Measure	Active Share		Diff.	Number of Positions		Diff.	Total
	Low	High	(*t*-test)	Concentrated	Diversified	(*P*-Value)	
Momentum							
N	476	1,042		754	783		3,031
Mean (BP)	24	−45	0.08	−88	7	0.01	−69
10th % (BP)	−590	−900		−980	−580		−800
Beat (%)	45	47	0.08	42	50	0.00	
% Beat by 1%	36	41		35	39		
% Miss by 1%	46	46		51	40		
Information Ratio							
N	600	918		754	840		2,983
Mean (BP)	8	−24	0.39	−90	7	0.00	−59
10th % (BP)	−500	−840		−810	−530		−750
Beat (%)	44	48	0.12	43	50	0.00	
% Beat by 1%	33	42		35	38		
% Miss by 1%	46	45		50	40		
***FF* Intercept *t*-Statistic**							
N	511	969		791	740		3,055
Mean (BP)	−90	−2	0.01	−94	−13	0.02	−56
10th %	−550	−890		−790	−600		−750
Beat (%)	35	47	0.00	39	46	0.01	
% Beat by 1%	25	41		32	36		
% Miss by 1%	52	45		53	42		

Source: CRSP Survivor-Bias-Free US Mutual Fund database, December 2014, http://www.petajisto. net/data.html, Nomura Securities, http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library. html.

Note: In this table we split fund-observations that belong to the top quartiles of momentum, information ratio, and FF intercept *t*-statistics in two ways: first, into those funds that are also in the top quartile of active share and those that are in the bottom active share quartile, and second, into those funds that are also in the top quartile of number of positions and those that are at the bottom quartile. *Momentum*: Cumulative return over the past 36 months minus the cumulative return of the benchmark over the same period. *Information Ratio*: Information ratio is defined as the mean of the differences in the monthly log returns between the fund return and the benchmark return over the prior 36 months, scaled by the tracking error. *FF Intercept t-Statistics*: For each fund, we regress the monthly returns of the fund, in excess of the risk-free rate, on the FF five-factor returns over the prior 36 months. The intercept of this regression measures the unique contribution of the fund manager beyond known risk factors. The *t*-statistic of this intercept is a measure of significant ability to generate returns that are not explainable by the traditional FF factors. *Active share* (%): Active share of a fund relative to its benchmark index as of each report date; obtained from Petajisto or Nomura. *Number of Positions*: The number of holdings in a fund; obtained from the CRSP mutual fund data. Sample observations are for funds that have 36 months of returns, active share, and number of positions during the period 2002–2013. Diff. (t-test) provides the significance level for tests that the mean excess returns is the same for both sub-groups, e.g. high and low quartiles.

quartile. These results indicate the superiority of holding a diversified portfolio.

The picture for top IR funds is similar to the picture for top momentum funds. Focusing first on the intersection of top IR funds with high and low active share, we see that the intersection of the top IR funds with the top quartile of active share have average subsequent one-year excess returns of −24 BP, compared to a positive average of 8 BP for the intersection with the low active share funds (with the differences not statistically significant).

However, if we look at the intersection of top IR funds with the top quartile of diversified funds (based on holdings), these funds have a **positive** return of 7 BP, compared to the intersection of top IR with concentrated funds, which have a **negative** return of −90 BP (with the differences statistically significant). The diversified top IR funds outperform their benchmark in the subsequent year 50% of the time, whereas the concentrated top IR funds only outperform in the subsequent year 43% of the time (with the difference statistically significant).

What about the top quartile of funds based on the *t*-statistic of the intercept in a Fama–French regression of a fund's return on the five Fama–French factors? In this case we see that if we intersect these top *t*-statistic funds with high active share funds, the average next-year return is −2 BP, compared to −90 BP for the intersection of top *t*-statistic funds with low active share. This result demonstrates that active share distinguishes among fund managers who are able to obtain a significant alpha from sources that are not related to the five Fama–French broad factors. High active share, in addition to manager skill as defined by high *t*-statistic, leads to superior performance.

Similar to the results for active share, we find that diversification is able to distinguish among funds in the top quartile of the *t*-statistic of the Fama–French intercept. The intersection of diversified funds with top *t*-statistic funds has an average next-year return of −13 BP, significantly better than their concentrated peers. Furthermore, as we saw in Table 2, drawdown for the diversified funds is −600 BP, lower not only than their concentrated peers, but also lower than the comparable group with high active share.

The results in Table 3 show the advantages of (i) high diversification in the case of top momentum and top IR funds, and (ii) high active share for top skill, as measured by the *t*-statistic of the five-factor Fame–French regression intercept.

In Table 4 we compare the subsequent one-year performance of diversified top IR funds and high active share, top skill (high *t*-statistic) funds. To ensure a full separation of the two groups, we eliminate the funds that belong to both groups.[10] We find both groups deliver similar, slightly positive subsequent one year returns: 1 BP for the high active share, top skill group, and 11 BP for the diversified, top IR group. Also, 51% of the diversified, top IR funds beat the benchmark in the subsequent year, compared to 48% for the high active share, top skill funds. However, the differences in next-year performance between the two groups are not statistically significant.

Correlation: The notion of diversification and concentration naturally leads to the question of correlation. While it may be intuitive to some investors that a single, concentrated fund is perhaps not entirely appropriate for their needs, a handful of such concentrated funds, if they are relatively uncorrelated, is likely to be desirable. A natural prior is that the correlation effects are more pronounced for concentrated funds than for diversified ones.

To address this issue, we again focus on diversified, top IR funds, and high active share, top skill

Table 4 One-year performance for funds in top quartile FF intercept's *t*-statistic/active share and top quartile IR/diversified.

	Top: FF Inter. *t*-stat. & Active Share	Top: IR & Diversified	Diff: (*P*-Value)
Panel A: All funds			
N	932	803	
Mean (BP)	1	11	0.76
Beat (%)	48	51	0.21
% Beat by 1%	42	38	
% Miss by 1%	45	39	
Panel B: Most uncorrelated funds			
N	12	12	
Mean (BP)	−8	36	0.76
Beat (%)	58	42	0.44

Source: CRSP Survivor-Bias-Free US Mutual Fund database, December 2014, http://www.petajisto.net/data.html, Nomura Securities, http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

Note: This table shows the one-year performance of two groups of funds. The first group is funds that have both a top quartile F − F intercept's *t*-statistic and active share. The second group is those funds that are in the top quartile of both IR and portfolio diversification. To ensure a full separation of the two groups, we eliminate the funds that belong to both. Panel A is for the full sample, while Panel B compares the performance of these two groups when five funds are selected each year from each group to have the lowest correlations among them. We begin this process by selecting the pair of funds that have the lowest positive (or highest negative) correlation from among all group members. We then select the fund which has the smallest sum of correlations with either of the funds in the selected pair. We repeat this process until we identify five funds each year, and examine the average performance in the subsequent year. *Information Ratio* (*IR*): Information ratio is defined as the mean of the differences in the monthly log returns between the fund return and the benchmark return over the prior 36 months, scaled by the tracking error. *FF Intercept t-Statistics*: For each fund, we regress the monthly returns of the fund, in excess of the risk-free rate, on the FF five-factor returns over the prior 36 months. The intercept of this regression measures the unique contribution of the fund manager beyond known risk factors. The *t*-statistic of this intercept is a measure of significant ability to generate returns that are not explainable by the traditional FF factors. *Active share* (%): Active share of a fund relative to its benchmark index as of each report date; obtained from Petajisto or Nomura. *Diversified*: Top quartile based on the number of positions in a fund; obtained from the CRSP mutual fund data. The data for this table is from all funds that had 36 months of returns, active share, and number of portfolio positions during the years 2002–2013. Diff: *t*-test is the significance level of a *t*-test that the two groups have equal means.

funds. Panel B of Table 4 compares the performance of these two groups when five funds are selected each year from each group to have the lowest correlations among them. This is intended to address a selection process whereby a consultant or an advisor selects five funds that have different characteristics such as core, value, growth, small, and mid-cap, or fundamental managers with different methodologies and expertise. We begin the process by selecting the pair of funds that have the lowest positive (or highest negative) correlation from among all group members. We then select the fund which has the smallest sum of correlations with either of the funds in the selected pair. We repeat this process until we identify five funds each year. We then examine

the average performance in the subsequent year.

Panel B illustrates that selecting uncorrelated, diversified top IR funds yields an average annual return of 36 BP. However, selecting uncorrelated funds from high active share, top skill (i.e. high t-statistic) funds yields an average return of −8 BP per year. While these differences are not statistically significant, the results are surprising in light of the intuitive prior that correlation matters more for concentrated than for diversified funds. This prior seems not to be supported by the data.

Summing up the results so far, we find that a statistically significant improvement in performance can be obtained by selecting top funds on either past performance (momentum) or IR, that are also the most diversified. We also find that selecting top skill funds (i.e. high t-statistics on five-factor Fama–French alpha) that also have high active share also leads to statistically significant performance improvement. However, these latter funds tend to have a greater drawdown than the diversified, top IR funds.

While the results for the diversified, top IR funds are compelling, we are still left with the issue that high IR can be achieved with a low alpha, but an even lower tracking error. We thus turn to our proposed measure of modified IR, which incorporates the investor's desired return in excess of the benchmark. We show that the Modified IR carries all of the same benefits as the standard IR, but it allows the investor, *ex ante*, to express a preference for return and to therefore orient their fund selection accordingly.

4 Modified IR

In this section of the paper we focus on Modified IR. We begin by introducing the metric and describing some of its basic characteristics.

4.1 Basic characteristics

The conventional IR is alpha (α) divided by tracking error (σ), i.e.

$$IR = \frac{\alpha}{\sigma}.$$

This measure is attractive because it is simple and addresses the issue of skill. It is not enough to simply select the managers with the highest alpha. High alphas are more likely if tracking error is high. But high tracking error can also lead to low alphas in the absence of skill. So, is past high alpha due to skill or luck? The IR measure addresses this question. The problem, however, is that high IR can be associated with low alpha if tracking error is low enough. And you can not pay your mortgage, or pension benefits for that matter, with a high IR; you need returns as well.

To address this issue, we preserve the attractive features of IR but modify it by subtracting a desired alpha (α^*) from the numerator. The Modified IR has the form

$$IR^* = \frac{\alpha - \alpha^*}{\sigma}.$$

We assume that investors are return seeking, so we restrict the desired alpha to be non-negative, i.e. $\alpha^* \geq 0$. The adjustment of alpha by subtracting the desired alpha seems innocent enough, after all we are just shifting alpha by a constant. But this seemingly innocuous transformation carries a punch. To see this, notice that the Modified IR is the conventional IR minus the desired alpha divided by tracking error. So for any positive desired alpha, dividing desired alpha by tracking error means that the Modified IR penalizes low tracking error relative to high tracking error. This feature is Modified IR's driving force.

More formally, if there are two funds with the same IR, the fund with the higher tracking error always has a higher Modified IR. To see this, let α_i and σ_i represent the alpha and tracking error

($i = [1, 2]$) of each fund. If the IR of each fund is equal, i.e. $\frac{\alpha_1}{\sigma_1} = \frac{\alpha_2}{\sigma_2}$, then

$$IR_2^* - IR_1^* = \alpha^* \left(\frac{\sigma_2 - \sigma_1}{\sigma_1 \sigma_2} \right).$$

This quantity is always positive as long as $\sigma_2 > \sigma_1$. Thus, holding IR constant, the Modified IR is a measure that discriminates *against* low tracking error funds for any positive desired alpha.

We can also solve for the desired alpha that makes an investor indifferent between choosing between a high IR fund and a lower IR one. In general, we have

$$IR_2^* - IR_1^* = IR_2 - IR_1 + \alpha^* \left(\frac{\sigma_2 - \sigma_1}{\sigma_1 \sigma_2} \right).$$

If fund 2 has a larger IR than fund 1, i.e. $IR_2 > IR_1$, then the desired alpha that makes an investor indifferent, on a Modified IR basis, between the two funds is given by

$$\alpha^* = (IR_2 - IR_1) \left(\frac{\sigma_1 \sigma_2}{\sigma_1 - \sigma_2} \right).$$

But since $IR_2 > IR_1$, requiring positive α^* means that $\sigma_1 > \sigma_2$. This means that a fund with lower IR but higher tracking error is more attractive than a fund with a higher IR but lower tracking error, as long as the desired alpha exceeds the value given in the above equation. This expression has the intuitive behavior that, holding the tracking error of the two funds constant, a larger difference in IRs requires a larger desired alpha. Furthermore, holding the difference in IRs constant, a smaller difference in the tracking error of the funds requires a larger desired alpha. This result also means that there is always a desired alpha that renders fund 1 more attractive, on a Modified IR basis, to fund 2, even though fund 2 has a higher IR.

Another way to think about Modified IR is in terms of portfolio optimization. The Modified IR is analogous to the Sharpe ratio if we associate alpha with expected return, desired alpha with the risk-free rate, and tracking error with standard deviation of returns (volatility). From an optimization perspective, higher risk-free rates lead to risky tangent portfolios with higher volatility. This is analogous to higher desired alpha leading to a preference for higher tracking error. This analogy also highlights the key point that desired alpha is not an optimization variable. Rather, it represents the investor preference for return over a benchmark.

This notion of investor preference as return over benchmark is also related to Roy's Safety First criteria (Roy 1952). In Roy's approach, the investor's goal is to minimize the probability that return falls short of the desired return. For normally distributed returns, minimizing the disaster probability is equivalent to maximizing the equivalent of the Modified IR. Analogously, for normally distributed alphas, maximizing the Modified IR is equivalent to minimizing the probability that alpha falls short of the desired level.

This notion of preference for return over a benchmark also highlights a feature of index funds, i.e. passive funds that attempt to replicate the return of a well-defined index portfolio. Index funds are attractive to investors because they are low cost. But their cost is not zero. Even if their tracking error is zero, as long as there is a cost to this passive investment the IR for an index fund is negative infinity. Thus, for investors with even a zero desired alpha, *ex ante*, index funds are not the appropriate *theoretical* choice. But the argument for indexing is not a theoretical one. It is based on the historical evidence, which we replicate here, that the typical fund underperforms its index. However, we provide evidence (Tables 6 and 7) that selection criteria based on high Modified IR in conjunction with adequate diversification have historically generated positive excess returns.

Table 5 Characteristics of top IR quintile funds for groups based on desired alpha levels.

Desired alpha (BP/year)	0	75	150	300	450	600	750	900
N	2,196	2,164	2,044	1,554	983	605	385	250
Expense Ratio (BP)	114	116	117	120	123	125	128	129
TNA	1,860	1,815	1,846	1,865	1,972	2,198	2,369	1,745
Turnover Ratio	72%	72%	73%	74%	77%	81%	84%	90%
% Sign. FF Intercept	9.10%	9.30%	9.80%	11.50%	14.30%	16.20%	18.50%	21.70%
% Beating target	46.20%	38.90%	33.20%	24.30%	17.70%	14.50%	15.90%	12.40%
Annual Exc. Return (BP)	−52	−59	−65	−69	−68	−69	−80	−91
No. of Holdings	132	126	121	113	110	107	105	102
Tracking Error (BP/year)	164	174	182	200	224	247	271	291

Source: CRSP Survivor-Bias-Free US Mutual Fund database, December 2014, http://www.petajisto.net/data.html, Nomura Securities, http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Note: This table reports the characteristics of funds that were in the top IR quintile in the past three years and also had returns in excess of the benchmark **and** the desired alpha during the prior three years for various desired alpha levels. The sample includes all funds during the years 2002–2013 with active share, and number of positions in the CRSP database. *Desired alpha (BP/year)*: The desired alpha is chosen by the investor. It represents the desired return of the fund beyond its benchmark. Desired alpha of zero means that the investor wishes just to obtain the return of the benchmark, e.g. a passive investor who invests in an index fund. *Expense Ratio* (*BP*): Funds' expense ratio measured in basis points obtained from the CRSP mutual fund data. *TNA* (*$ Million*): Funds' total net assets obtained from the CRSP mutual fund data. *Turnover Ratio* (%): Funds' turnover ratio obtained from the CRSP mutual fund data. *% Sign. FF Intercept*: Percentage of funds with significant intercept (at the 5% two-sided level) in a regression of the fund return (minus the risk-free rate) on the five FF factors during the prior 36 months. *% Beating target*: Percentage of funds that actually beat their target in the subsequent year. *Annual Exc. Return* (*BP*): The fund's return during the subsequent year measured in basis points in excess of the benchmark, i.e. the cumulative return on the fund in the next 12 months minus the cumulative return on the benchmark at the same time period. *No. of Holdings*: The fund's average number of holdings from the CRSP database.

N (*for Holdings*)*:* The number of observations with available holdings (2002–2013). *Tracking Error* (*BP/year*): The tracking error is defined as the standard error of the differences between the fund and the benchmark log monthly returns over the prior 36 months.

Desired alpha and fund characteristics

Table 5 reports the characteristics of funds that were in the top Modified IR quintile in the past three years and were also able to have returns that were in excess of the benchmark **and** the desired alpha during the prior three years for various assumed desired alpha levels.[11] As is intuitively reasonable, higher desired alpha leads to fewer funds that both exceeded the desired alpha and are in the top IR quintile. Skipping to the last line of Table 5, we see that, consistent with our prior assertion, choosing a higher desired alpha is tantamount to choosing funds with higher tracking error. Also noticeable is that in order to achieve the higher desired alpha in the past three years, funds have to turn over their portfolios faster, be more concentrated in terms of the number of positions in their portfolio, and typically be compensated for that by higher expense ratios. They also are more likely to have a significant intercept (at the 5% two-sided level) in the five-factor Fama–French regression of the fund returns on the five-factors.

However, a closer look at Table 5 indicates an interesting finding about future returns. In particular, the realized return beyond the benchmark on the funds during the subsequent year is around −60 BP for low levels of desired alpha, and surprisingly is not better with funds that were selected because they delivered a higher desired

alpha in the prior three years. Similarly, when we examine the percentage of funds that were able to exceed the desired alpha in the subsequent year, the percentage drops monotonically from 46% with a desired alpha of zero to only 12.4% for funds that were selected because they were able to exceed a desired alpha of 900 BP in the prior three years. Thus, while relying on the Modified IR to select funds is a good predictor of future performance, one should realize that simply selecting top Modified IR funds that were historically able to exceed their desired alpha in the past three years does not necessarily guarantee that these funds will beat their benchmark in the future.

It is worthwhile to reiterate our earlier point about the desired alpha and tracking error. Selecting a higher desired alpha means that funds which deliver this desired alpha are *ex ante* expected to also have higher tracking error. With a higher past tracking error there is a higher likelihood that their future performance will be more erratic and not exceed the desired alpha, as we observe in Table 5. The question is whether we can enhance the fund selection in a way that increases the likelihood that future returns will be more consistent with past performance. To do that, we focus on funds that have greater breadth in terms of number of positions they hold. We conjecture that funds with fewer positions take more idiosyncratic risk and that this increased idiosyncratic risk makes outperformance less likely to be repeated in the future. To examine this conjecture, we split funds in the top Modified IR quintile that managed to exceed the desired alpha in the prior three years into those with fewer than 100 holdings, and those with more than 100 holdings.[12] The evidence for these two types of funds is available in Table 6.

As can be seen in Table 6, as desired alpha increases, the number of concentrated funds that meet the selection criteria (top quartile of IR and exceeding desired alpha) is greater than the number of diversified funds that meet the same criteria. While for a desired alpha of 0 BP there are 1,374 fund-observations that satisfy the selection criteria with less than 100 positions, there are 822 such fund-observations with over 100 positions. However, when we require a desired alpha of 300 BP, the numbers drop to 1,077 and 477, respectively. At 900 BP, there are hardly any observations for the funds with more than 100 positions. What can we say about the characteristics of the two classes of funds? The more concentrated funds have higher expense ratios, are smaller, have lower turnover, have higher tracking error in the prior three years, and have been less likely to have a significant (at the 5% two-sided level) intercept in the five-factor Fama–French regression during the prior three years than the more diversified funds. However, when we examine their realized returns in the subsequent year, the diversified funds have positive excess returns whereas the concentrated funds have negative returns. Also, the diversified funds are able to exceed the benchmark with a higher frequency in the following year. Thus, more diversified funds seem to outperform the concentrated funds in the same category (top IR quintile and returns that exceeded the target in the past). Note that, almost by construction, the diversified funds have lower active share than the more concentrated funds. Thus, higher active share is not necessarily a desirable quality for a specific category.

Table 7 further examines the differences between concentrated and diversified funds in a manner that is similar to an investor who wishes to select a small subset of funds for investment. This is a more realistic setting for financial advisors, for defined benefit or defined contribution plans in fund selection, or what consultants go through in manager searches. Each year, we randomly select five funds out of the pool of funds that are in the top quintile of Modified IR and that exceeded

Table 6 Characteristics of top IR quintile funds for groups based on desired alpha levels and the number of holdings.

Desired alpha (BP/year)	0	75	150	300	450	600	750	900
Panel A: Funds with less than 100 holdings								
Annual Exc. Return (BP)	−95	−95	−103	−105	−116	−120	−159	−193
N	1,374	1,413	1,381	1,077	697	433	280	183
Expense Ratio (BP)	119	121	121	124	126	128	130	132
TNA	984	994	985	981	979	1,025	1,099	876
Turnover Ratio	67%	68%	69%	70%	72%	76%	79%	87%
% Sign. FF Intercept	8.28%	8.20%	8.43%	10.00%	11.77%	13.39%	15.92%	18.42%
% Beating target	43.53%	38.00%	32.21%	24.07%	17.04%	12.95%	9.34%	4.21%
No. of Holdings	57	57	56	55	54	52	50	50
Active Share	84%	85%	86%	89%	88%	89%	90%	90%
Tracking Error (BP/year)	181	190	195	212	236	256	280	299
Panel B: Funds with more than 100 holdings								
Annual Exc. Return (BP)	16	5	8	10	39	48	109	160
N	822	751	663	477	286	172	105	67
Expense Ratio (BP)	104	106	108	112	116	116	12	122
TNA	3,229	3,255	3,518	3,744	4,199	4,908	5,403	3,890
Turnover Ratio	80%	80%	82%	85%	89%	94%	100%	98%
% Sign. FF Intercept	10.37%	11.11%	12.38%	14.66%	19.88%	22.68%	24.79%	29.87%
% Beating target	50.27%	40.59%	35.02%	24.81%	19.26%	18.04%	17.36%	14.29%
No. of Holdings	248	247	247	235	235	233	237	231
Active Share	73%	75%	76%	79%	77%	78%	78%	77%
Tracking Error (BP/year)	137	146	156	176	198	224	250	271

Source: CRSP Survivor-Bias-Free US Mutual Fund database, December 2014, http://www.petajisto.net/data.html, Nomura Securities, http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

Note: This table reports the characteristics of funds that were in the top IR quintile in the past three years and also had returns in excess of the benchmark **and** the desired alpha during the prior three years for various desired alpha levels. Panel A reports the characteristics of funds with less than 100 holdings, whereas Panel B reports the characteristics of funds with more than 100 holdings. *Desired alpha* (*BP/year*): The desired alpha is chosen by the investor. It represents the desired return of the fund beyond its benchmark. Desired alpha of zero means that the investor wishes just to obtain the return of the benchmark, e.g. a passive investor who invests in an index fund. *Expense Ratio* (*BP*): Funds' expense ratio measured in basis points obtained from the CRSP mutual fund data. *TNA* (*$ Million*): Funds' total net assets obtained from the CRSP mutual fund data. *Turnover Ratio* (%): Funds' turnover ratio obtained from the CRSP mutual fund data. *% Sign. FF Intercept*: Percentage of funds with significant intercept (at the 5% two-sided level) in a regression of the fund return (minus the risk-free rate) on the five FF factors during the prior 36 months. *% Beating Target*: Percentage of funds that actually beat their target in the subsequent year.

Annual Exc. Return (*BP*): The fund's return during the subsequent year measured in basis points in excess of the benchmark, i.e. the cumulative return on the fund in the next 12 months minus the cumulative return on the benchmark at the same time period. *No. of Holdings*: The fund's average number of holdings from the CRSP database *N* (*for Holdings*): The number of observations with available holdings (2002–2013).

Active share: Active share of a fund relative to its official benchmark index as of each report date; obtained from Petajisto and Nomura.

Tracking Error (*BP/year*): The tracking error is defined as the standard error of the differences between the fund and the benchmark log monthly returns over the prior 36 months.

Table 7 Simulation results: Future returns of concentrated vs. diversified funds.

Desired alpha		Number of holdings		Difference	Significance (*P*-Value)	No. of years
		<100	≥100			
0 BP	Mean Annual Return	−104	−32	−72	0.0001	12
	Median Annual Return	−111	−36	−75	0.0001	12
	St. Dev. Annual Return	589	477	112	0.0001	12
	% Beat	42%	48%	−6%	0.0001	12
75 BP	Mean Annual Return	−109	−24	−85	0.0001	12
	Median Annual Return	−122	−34	−88	0.0001	12
	St. Dev. Annual Return	611	495	116	0.0001	12
	% Beat	37%	40%	−3%	0.0001	12
150 BP	Mean Annual Return	−104	−16	−88	0.0001	12
	Median Annual Return	−119	−24	−95	0.0001	12
	St. Dev. Annual Return	623	518	105	0.0001	12
	% Beat	32%	36%	−4%	0.0001	12
300 BP	Mean Annual Return	−113	15	−128	0.0001	12
	Median Annual Return	−124	1	−125	0.0001	12
	St. Dev. Annual Return	661	573	88	0.0001	12
	% Beat	25%	27%	−2%	0.0001	12
450 BP	Mean Annual Return	−102	82	−184	0.0001	12
	Median Annual Return	−102	48	−150	0.0001	12
	St. Dev. Annual Return	689	699	−10	0.0470	12
	% Beat	20%	23%	−3%	0.0001	12
600 BP	Mean Annual Return	−111	93	−204	0.0001	9
	Median Annual Return	−107	126	−233	0.0001	9
	St. Dev. Annual Return	697	581	116	0.0001	9
	% Beat	17%	21%	−4%	0.0001	9
750 BP	Mean Annual Return	−207	31	−238	0.0001	5
	Median Annual Return	−157	90	−247	0.0001	5
	St. Dev. Annual Return	708	604	104	0.0001	5
	% Beat	7%	13%	−6%	0.0001	5

Source: CRSP Survivor-Bias-Free US Mutual Fund database, December 2014, http://www.petajisto.net/data.html, Nomura Securities, http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

Note: This table reports the future annual returns (BP) of diversified vs. concentrated funds generated in the following simulation. Each year, we randomly select five funds out of the pool of funds that are in the top quintile of Modified IR and also exceeded the desired alpha in the prior three years. We repeat the process 1,000 times, separately for the concentrated and the diversified funds, as long as there are at least 10 funds to select from in each category. The table reports the mean and median fund return minus the benchmark return in the subsequent 12 months, as well as the standard deviation. It is based on data for 2002–2013. %Beat is the percentage of funds that exceeded the desired alpha in the subsequent year. For desired alpha of 600 BP and higher there were fewer than 10 funds in either the concentrated or diversified funds category, so fewer than 12 yearly simulations are reported. The table reports the significance levels of a matched pair *t*-test (or Wilcoxon test for the median) for the differences between the selected concentrated and diversified funds.

Table 7 (*Continued*)

Desired alpha		Number of holdings		Difference	Significance (*P*-Value)	No. of years
		<100	≥100			
900 BP	Mean Annual Return	−265	38	−303	0.0001	4
	Median Annual Return	−217	44	−261	0.0001	4
	St. Dev. Annual Return	634	578	56	0.0001	4
	% Beat	4%	4%	−1%	0.0073	4

the desired alpha in the prior three years. We repeat the process 1,000 times, separately for the concentrated and the diversified funds, as long as there are at least 10 funds to select from in each category. This is similar to the procedure used by Domian and Nanigian (2014). We then examine the future performance in the subsequent year. Table 7 contains statistics about the future returns in this simulation, as well as statistical tests of the differences between the concentrated and diversified funds.

As can be seen in Table 7, the diversified funds have higher excess returns in the subsequent year than the concentrated funds with statistically significant differences for all levels of desired alpha. They also have lower standard deviation of the future excess returns for all desired alpha levels, except for alpha of 450 BP. Finally, they are more likely to exceed the desired alpha in the subsequent year than the concentrated funds, and these differences are also statistically significant. The reason for the decline in the number of years for desired alpha levels above 600 BP is that there are fewer than 10 funds in the diversified category in some years that passed the selection criteria.

One of the main takeaways from Tables 6 and 7 is that selection of funds based on particular combinations of Modified IR and diversification has historically generated positive excess returns for many desired levels of alpha.

Robustness tests

To test whether the Global Financial Crisis (GFC) has affected our results because the modified IR positioned funds that controlled their risk exposures in the top quintile, we removed the years 2008–2010 from our dataset, and repeated all the analyses. We find that "winner" funds that were in the top modified IR quintile and that were able to beat the desired alpha in the prior 36 months, had a significantly higher future return if they were more diversified (held more than 100 positions in their portfolio) than funds which were concentrated. Thus, the GFC is not the cause of our results.

We also repeated our tests in three sub-periods, up to 1999, 2000–2007, and 2008–2013, each including about one third of the sample observations. Because the number of fund holdings is available in the CRSP database only from 2002, the results for the first sub-period are mostly about the superiority of IR and the characteristics of "winner" funds according to the modified IR criterion. We find that in the first sub-period, which included the Internet bubble, momentum had a slightly higher return than IR, but IR was superior in terms of the standard deviation of excess returns. Other fund characteristics are similar in the first sub-period to the entire sample period.

In the second and third sub-periods, the results are largely similar to the results of the entire sample. IR dominates other predictors. "Winner" funds

with greater desired alpha have similar characteristics to those we find in the entire sample. We found that "winner" funds that were diversified had significantly higher excess future returns than "winner" funds with fewer than 100 positions, consistent with the results for the entire sample.

We therefore conclude that our results are robust to different time periods.

Replication of key results using eVestment alliance data

As an additional robustness check, we used performance data on institutional products from eVestment Alliance to replicate the key results in this study. For this analysis we use gross returns on long-only active US equity products spanning different styles (core, growth and value), investment approaches (fundamental, quantitative and combined) and market cap size segments (all cap, large-cap, mid-cap, smid-cap, small-cap) beginning in 1994. Because we use 36 months of return data for construction, our testing begins from the end of 1996.

From the eVestment Alliance data we confirm that IR seems to have the strongest predictive power of future returns. When we examine the characteristics of funds that were in the top IR quintile in the past three years and were also able to have returns that were in excess of the benchmark **and** the desired alpha during the prior three years for various desired alpha levels, we find that again choosing a higher desired alpha is equivalent to choosing funds with higher tracking error. Similar to the findings on mutual funds, in order to achieve the higher desired alpha in the past three years, institutional investors had to turn over their portfolios faster and to be more concentrated in terms of the number of positions in their portfolio. They also were able to more often have a past significant intercept in the Fama–French regression of the fund returns on the five-factors during the prior 36 months.

One of our key findings using mutual fund data is that the number of holdings is a useful measure to predict consistency of future returns—funds with larger numbers of holdings have a higher likelihood of future returns that are consistent with past performance. In this context, what can we say about the characteristics of the concentrated versus diversified products that met the selection criteria? Similar to the mutual fund findings, the more concentrated funds are smaller, have lower turnover, have higher tracking error in the prior three years, and have been more likely to have a significant intercept in the five-factor Fama–French regression during the prior three years than the more diversified funds. Further, when we examine their realized returns in the subsequent year, the diversified products have actually outperformed the concentrated products.

5 Summary and conclusions

In this paper we study the future performance of various measures for the selection of funds. We show that there are two paths that are likely to yield a portfolio of funds with superior future performance. In the first approach, we select top funds based on either past performance (i.e. momentum) or information ratio, and then keep only those top funds that are also highly diversified (as measured by the number of holdings). The second approach is to identify manager skill via the regression of fund returns on the five Fama–French factors (the *t*-statistic of the intercept), and then keep only those funds with high active share. Both of these approaches identify funds that outperform other top funds within their respective categories (top IR or manager's skill). We also find that the approach based on diversified top IR funds yields slightly higher future excess returns and lower potential drawdown,

which we attribute to higher consistency of returns.

To address the issue that high IR can be achieved with low alpha but even lower tracking error, we introduce a new measure—the Modified IR. Instead of scaling alpha by tracking error, as is done in the conventional IR calculation, the Modified IR scales alpha *minus an investor-specified desired alpha*. This seemingly simple modification effectively deals with the problem of high IR but low return by requiring that the investors specify their return preferences via the desired alpha. When we examine the properties of our Modified IR measure we find that as the desired alpha increases, those funds in the top quintile during the past 36 months according to the Modified IR have a higher tracking error, fewer stock holdings, higher expense ratios, higher active share, and higher turnover ratios. These attributes are generally consistent with what one would expect is required to achieve a higher alpha level. However, we also find that the subsequent performance of funds that rank well on the Modified IR and that exceed the investor's desired alpha in the past 36 months do not necessarily increase with the desired alpha levels. This result highlights the boiler-plate, but nevertheless important, risk mitigation clause of "past performance is not a guarantee of future performance". However, more diversified funds have demonstrated a better chance of repeating their prior outperformance.

When we combine diversification with Modified IR in the same way that we do for the conventional IR, we find similar results. In particular, we find that funds in the top quintile of Modified IR with a larger number of positions are more likely to have better future performance than those that are more concentrated.[13]

The Modified IR suggests a reframing of the fund selection problem. First, investors need to specify their investment objective—what alpha do they desire, or need? Second, investors need to identify the funds with the highest likelihood of achieving that objective. We show that, using historical fund performance data, that high Modified IR with more diversification is an effective solution.

Appendix A: Short literature review

There is an extensive literature about methods or measures that can be used to predict future performance of funds. In this short review, we shall focus on the more recent approaches, and on those studies that focus on IR. The persistence of fund returns was documented in early studies by Kahn and Rudd (1995), Carhart (1997) and Wermers (1997). On the tracking error front, Israelsen and Cogswell (2007) note that actively managed mutual funds with low tracking error exhibit lower alpha, higher beta, and lower average performance compared to funds with high tracking error. However, they also examine information ratio as a better way to evaluate funds. High IR funds demonstrate higher alpha, lower beta, and higher returns than funds with low IR. Additionally, they find that high IR funds demonstrate significantly higher tracking error than low IR funds. Bollen and Busse (2005) show that there appears to be persistence in fund performance in the short-run, but not with respect to the long-run. After controlling for stock-level momentum and other stock attributes, there does not appear to be evidence of mutual fund return persistence.

Few papers have directly examined the relationship between information ratios and future fund performance. Gupta *et al.* (1999) examine the relation between IR and fund performance. Their main conclusion, consistent with ours, is that it is the information ratio not the alpha and/or tracking error that is the strongest predictor of persistence of manager performance. Bossert *et al.* (2010) also examines whether IR is a useful and reliable performance measure to evaluate mutual fund managers. They find that IR varies over time

and also across different fund categories, and that the quality and reliability of the IR is dependent on certain estimation choices (e.g. a long-term track record is required to be able to separate lucky managers from skilled ones).

While the IR is a powerful tool to assess the skill set of an active manager, its usefulness can be significantly affected by the construction methods and the characteristics and statistical properties returns. Goodwin (1998) shows that the choice of benchmark to match the style of the manager can make a dramatic difference. In addition to the IR being highly sensitive to the choice of benchmark, it is also subject to substantial estimation uncertainty and can be affected by the way the returns are annualized. Israelsen (2005) warns that the reliability of the IR decreases when the excess returns are negative: the use of IR might generate anomalous rankings. When fund returns are negative, the IR for the more volatile fund will be higher. Hübner (2007) compares the performance of the IR with the alpha and generalized Treynor ratio using a sample of directional mutual funds with different styles. He compares both the stability (robustness under different asset pricing models) and the precision (a good measure should be able to provide true ranking of funds based on investor preferences) of the measures. He shows that the IR displays the poorest level of precision and stability. He attributes the poor empirical results for the IR to its heavy reliance on variance as the measure of risk. Measurement issues due to a significantly non-normal distribution of regression residual result in the downward bias of the IR. Similarly, Nanigian (2011) expresses concern that the assumptions underlying the IR do not take into consideration the utility theory. The IR assumes that the investor utility does not vary with their level of wealth and that all investors have identical levels of risk aversion. He calls for a new, more informative measure of fund performance that incorporates the non-linear utility function and different levels of risk aversion across individuals. Constable and Armitage (2006) argue that while the IR is a commonly accepted measure of success of asset managers, it is just one aspect of manager success and does not really inform investors about how the manager achieved this ratio. When investors are provided with the IR at the end of an investment period, they cannot really assess the string of successes and failures that led to the outcome presented in the IR. The limitation of IR arises from its construction, as it captures information from the first two moments of the distribution of expected returns.

There are many papers that attempt to identify the 'activeness' of a fund and show how it can be used to predict future performance. Cremers and Petajisto's (2009) paper is the most notable paper in this area. They document the relation between active share and future fund performance for US funds; and show high active share funds with higher tracking error generate stronger returns. Cremers *et al.* (2016) extend this finding to international funds to show that active share is positively related to future benchmark adjusted returns and excess four-factor alphas.

Other support for better performance of concentrated portfolios comes from Kacperczyk *et al.* (2005), which show that funds with concentrated select industry holdings perform better, after controlling for risk and style differences using various performance measures. Baks *et al.* (2006) and Brands *et al.* (2005) show that portfolios with concentrated stock holdings outperform their diversified counterparts. Amihud and Goyenko (2013) propose that fund performance can be predicted by its *R*-squared, obtained from a regression of its returns on a multi-factor benchmark model. Lower *R*-squared is said to measure greater selectivity or active management. It is shown to significantly predict higher alpha. Related to Amihud and Goyenko (2013), Derwall and Huji (2011) in

a global setting identify high tracking error funds using the *R*-squared from a market model. They show these concentrated funds generate better performance. More interestingly, they show that the observed relation between portfolio concentration and performance is mostly driven by the breadth of the underlying fund strategies (country, sector and styles); not just by fund managers' willingness to take big bets.

Doshi *et al.* (2015) argue that active share and tracking error are sensitive to correctly identifying the funds' investment benchmarks or peer groups. They propose a new measure of activeness, the active weight, which is the absolute difference between the value weights and the actual weights held by a fund, averaged across its holdings. This measure is shown to be effective in predicting fund performance after controlling for activeness proxies such as industry concentration ratio, return gap, active share, and *R*-squared.

Fulkerson and Riley (2015) examine why funds with high active share outperform. They find that two-thirds of the outperformance can be attributed to the funds' ability to select out-of-benchmark stocks. In a similar spirit, Cremers and Pareek (2016) find that among high active share funds only those with patient investment strategies outperform their benchmarks.

Frazzini *et al.* (2015) are more critical of active share. Besides questioning the theoretical support for why active share should predict outperformance, they show that the results of Cremers and Petajisto are a product of benchmark effects. Namely, funds benchmarked to large-cap indexes typically have a lower active share, compared with funds benchmarked against small-cap indexes. When funds are ranked on active share within each benchmark size group, there is no significant difference in performance between Stock Pickers and Closet Indexers.

Notes

1 Appendix A contains a brief literature review.

2 We use only funds with CRSP Objective Codes of EDCL, EDCM, EDCS, EDYG, EDYB, which are the codes for Equity, Domestic, and then large-, mid-, and small-cap, growth and value, respectively. We use $15 million minimum size as in Petajisto (2013).

3 We balance the need for a longer time-series of returns to obtain better estimates of past alpha and standard error with a desire to keep economic, market and managerial conditions for the fund similar. We settle on 36 months in line with consultants' and investors' use of a three-year track record.

4 Petajisto's data are available at http://www.petajisto.net/data.html.

5 See Mezrich (2015) for a discussion of data sources and methodology.

6 The factor data are obtained from http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

7 Recall that the CRSP mutual funds database reports returns that are net of expenses.

8 The top quartile of active share has a median active share of 94%, as compared to 59% for the bottom quartile.

9 The top quartile of holdings has a median of 205 positions, as compared to 38 for the bottom quartile.

10 These are funds whose managers exploit idiosyncratic sources of alpha, have high active share, but also have many portfolio holdings. This is likely due to very small positions in many firms that are held as remnants of previously desired positions, or new positions that are slowly being built. Because the mean returns for the remaining funds have slightly higher returns than those reported in Table 3, the funds in the intersection of these two groups had negative return performance in the subsequent year.

11 It should be noted that the desired alpha is a decision parameter by the investor based on the investor needs for return levels, risk, and performance expectations for the benchmark. It is an *ex ante* parameter, but has implications for the type of funds that were *ex post* able to beat that desired alpha level in the past. Since we do not derive the desired alpha level from any theoretical model, we simulate it by specifying various levels of desired alpha.

12 Unlike earlier, when we used the cross-sectional distribution of funds every year to select the top quartile, we use here a fixed 100 positions as a cutoff. This has the advantage of providing a robustness check to our earlier results. The 100 positions threshold is at

roughly the 65% level on average, and is a reasonable proxy to distinguish quant from fundamental funds. This is confirmed separately in a different study using the eVestment data.

[13] It should be noted that our results are for the average diversified and concentrated funds. In each group there are funds that outperform or underperform not only their own group's averages, but also the other group's averages.

Acknowledgment

The authors acknowledge active share data obtained from Antti Petajisto's website which relates to his 2013 study. The authors also acknowledge more recent active share data that was provided by Joseph Mezrich at Nomura Securities. The paper benefited from comments received in presentations at the Northfield Information Services Annual Research Conference, at QMa's research meeting, at various units within Prudential Financial, at Deutsche Bank, at Barclays' Quant conference and at Nomura's Quant conference.

References

Amihud, Y. and Goyenko, R. (2013). "Mutual Fund R2 as Predictor of Performance," *Review of Financial Studies* **26**(3), 667–694.

Baks, K., Busse, J., and Green, C. (2006). "Fund Managers Who Take Big Bets: Skilled or Overconfident," Working Paper.

Bossert, T., Fuss, R., Rindler, P., and Schneider, C. (2010). "How "Informative" Is the Information Ratio for Evaluating Mutual Fund Managers?," *The Journal of Investing* **19**(1), 67–81.

Bollen, N. P. B. and Busse, J. A. (2005). "Short-Term Persistence in Mutual Fund Performance," *Review of Financial Studies* **18**(2), 569–597.

Brands, S., Brown, S., and Gallagher, D. (2005). "Portfolio Concentration and Investment Manager Performance," *International Review of Finance* **5**(3–4), 149–174.

Carhart, M. M. (1997). "On Persistence in Mutual Fund Performance," *Journal of Finance* **52**(1), 57–82.

Constable, N. and Armitage, J. (2006). "Information Ratios and Batting Averages," *Financial Analysts Journal* **62**(3), 24–31.

Cremers, K. J. M., Ferreira, M. A., Matos, P., and Stark, L. (2016). "Indexing and Active Fund Management: International evidence," *Journal of Financial Economics* **120**(3), 539–560.

Cremers, M. and Pareek, A. (2016). "Patient Capital Outperformance – The Investment Skill of High Active Share Managers Who Trade Infrequently," *Journal of Financial Economics* **122**(2), 288–306.

Cremers, K. J. M. and Petajisto, A. (2009). "How Active is Your Fund Manager? A New Measure that Predicts Performance," *Review of Financial Studies* **22**, 3329–3365.

Derwall, J. and Huij, J. (2011). "Global Equity Fund Performance, Portfolio Concentration, and the Fundamental Law of Active Management," *Journal of Banking and Finance* **35**, 155–165.

Domian, D. and Nanigian, D. (2014). "R^2 and the Benefits of Multiple-Fund Portfolios," *Journal of Investing Fall* 43–64.

Doshi, H., Elkamhi, R. and Simutin, M. (2015). "Managerial Activeness and Mutual Fund Performance," *Review of Asset Pricing Studies* **5**(2), 156–184.

Frazzini, A., Friedman, J., and Pomorski, L. (2015). "Deactivating Active Share," Working Paper.

Goodwin, T. H. (1998). "The Information Ratio," *Financial Analysts Journal* **54**(4), 34–43.

Gupta, F., Prajogi, R., and Stubbs, E. (1999). "The Information Ratio and Performance," *The Journal of Portfolio Management* **26**(1), 33–39.

Hübner, G. (2007). "How Do Performance Measures Perform?," *Journal of Portfolio Management* **33**(4), 64–74.

Israelsen, C. L. (2005). "A Refinement to the Sharpe Ratio and Information Ratio," *Journal of Asset Management* **5**(6), 423–427.

Israelsen, C. and Cogswell, G. (2007). "The Error of Tracking Error," *Journal of Asset Management* **7**, 419–424.

Kacperczyk, M., Sialm, C., and Zheng, L. (2005). "On the Industry Concentration of Actively Managed Equity Mutual Funds," *Journal of Finance* **60**, 1983–2011.

Kahn, R. N. and Rudd, A. (1995). "Does Historical Performance Predict Future Performance?," *Financial Analysts Journal* **51**(6), 43–52.

Mezrich, J. (2015). "Does High Active Share Help Active Managers?," *Nomura Quantitative Investment Strategy*.

Nanigian, D. (2011). "A More Informative Measure of Active Fund Performance," *Journal of Financial Planning* **24**(11), 54–60.

Petajisto, A. (2013). "Active Share and Mutual Fund Performance," *Financial Analysts Journal* **69**(4), 73–93.

Roy, A. D. (1952). "Safety First and the Holding of Assets," *Econometrica: Journal of the Econometric Society*, 431–449.

Wermers, R. (1997). "Momentum Investment Strategies of Mutual Funds, Performance Persistence and Survivorship Bias," Working paper.

Keywords: Fund performance; fund selection; fundamental-quantitative; diversification