How Active Is Your Fund Manager? A New Measure That Predicts Performance*

Martijn Cremers
International Center for Finance
Yale School of Management

Antti Petajisto†
International Center for Finance
Yale School of Management

October 3, 2007

Abstract

To quantify active portfolio management, we introduce a new measure we label Active Share. It describes the share of portfolio holdings that differ from the benchmark index. We determine the type of active management for a portfolio by measuring it in two dimensions using both Active Share and tracking error volatility. We apply this approach to the universe of all-equity mutual funds to characterize how much and what type of active management they practice. We test how active management is related to fund characteristics such as size, expenses, and turnover in the cross-section, and we examine the evolution of active management over time. Active management also predicts fund performance: funds with the highest Active Share significantly outperform their benchmark indexes both before and after expenses, and they exhibit strong performance persistence even after controlling for momentum. Non-index funds with the lowest Active Share underperform.

JEL classification: G10, G14, G20, G23

Keywords: Portfolio management, Active Share, tracking error, closet indexing

---

*We wish to thank Nick Barberis, Jonathan Berk, Lauren Cohen, Roger Edelen, Frank Fabozzi, Andrea Frazzini, Will Goetzmann, John Griffin, Owen Lamont, Juhani Linnainmaa, Ludovic Phalippou, Andrei Shleifer, Clemens Sialm, Matthew Spiegel, Lu Zheng, and Eric Zitzewitz for comments, as well as conference and seminar participants at the AFA 2007 Annual Meeting, CFEA 2006, CRSP Forum 2006, EFA 2007 Annual Meeting, FRA 2006 Annual Meeting, NBER Asset Pricing Meeting, Morningstar Investment Conference, Super Bowl of Indexing, Federal Reserve Bank of New York, Securities and Exchange Commission, Amsterdam University, Boston College, Helsinki School of Economics, ISCTE Business School (Lisbon), Tilburg University, University of Chicago, University of Maryland, University of Vienna, and Yale School of Management. We are also grateful to Barra, Frank Russell Co., Standard & Poor's, and Wilshire Associates for providing data for this study.

†Corresponding author: Antti Petajisto, Yale School of Management, P.O. Box 208200, New Haven, CT 06520-8200, tel. +1-203-436-0666, antti.petajisto@yale.edu.

1 Introduction

An active equity fund manager can attempt to outperform the fund's benchmark only by taking positions that are different from the benchmark. Fund holdings can differ from the benchmark holdings in two general ways: either because of stock selection or factor timing (or both).[1] Stock selection involves picking individual stocks that the manager expects to outperform their peers. Factor timing involves time-varying bets on systematic risk factors such as entire industries, sectors of the economy, or more generally any systematic risk relative to the benchmark index. Because many funds favor one approach over the other, it is not clear how to quantify active management across all funds.

Tracking error volatility (hereafter just "tracking error") is the traditional way to measure active management. It represents the volatility of the difference between a portfolio return and its benchmark index return. However, the two distinct approaches to active management contribute very differently to tracking error, despite the fact that *either* of them could produce a higher alpha.

For example, the T. Rowe Price Small Cap fund is a pure stock picker which hopes to generate alpha with its stock selection *within* industries, but it simultaneously aims for high diversification *across* industries. In contrast, the Morgan Stanley American Opportunities fund is a "sector rotator" which focuses on actively picking entire sectors and industries that outperform the broader market while holding mostly diversified (and thus passive) positions within those sectors. The tracking error of the diversified stock picker is substantially lower than that of the sector rotator, suggesting that the former is much less active. But this would be an incorrect conclusion – its tracking error is lower simply because individual stock picks allow for greater diversification, even while potentially contributing to a positive alpha.

Instead, we can compare the portfolio *holdings* of a fund to its benchmark index. When a fund overweights a stock relative to the index weight, it has an active long position in it, and when a fund underweights an index stock or does not buy it at all, it implicitly has an active short position in it. In particular, we can decompose any portfolio into a 100% position in its benchmark index plus a zero-net-investment long-short portfolio on top of that. For example, a fund might have 100% in the S&P 500 plus 40% in active long positions and 40% in active short positions.[2]

We propose the size of this active long-short portfolio (40% in the previous example)

[1] The basic idea has been presented and discussed by Fama (1972), Brinson, Hood, and Beebower (1986), Daniel, Grinblatt, Titman, and Wermers (1997), and many others.

[2] Asness (2004) discusses the same decomposition, albeit from the point of view of tracking error alone.

as a new measure of active management, and we label this measure the Active Share of a portfolio. Since mutual funds almost never take actual short positions, their Active Share will always be between zero and 100%. Active Share can thus be easily interpreted as the "fraction of the portfolio that is different from the benchmark index."

We argue that Active Share is useful for two main reasons. First, it provides information about a fund's potential for beating its benchmark index – after all, an active manager can only add value relative to the index by deviating from it. Some positive level of Active Share is therefore a necessary (albeit not sufficient) condition for outperforming the benchmark.

Second, while Active Share is a convenient stand-alone measure of active management, it can also be used together with tracking error for a more comprehensive picture of active management, allowing us to distinguish between stock selection and factor timing. The main conceptual difference between the measures is that tracking error incorporates the covariance matrix of returns and thus puts significantly more weight on correlated active bets, whereas Active Share puts equal weight on all active bets regardless of diversification. Hence, we can choose tracking error as a reasonable proxy for factor bets and Active Share for stock selection.[3]

Using these proxies, the two dimensions of active management are illustrated in Figure 1. A diversified stock picker can be very active despite its low tracking error, because its stock selection within industries can still lead to large deviations from the index portfolio. In contrast, a fund betting on systematic factors can generate a large tracking error even without large deviations from index holdings. A concentrated stock picker combines the two approaches, thus taking positions in individual stocks as well as in systematic factors. A "closet indexer" scores low on both dimensions of active management while still claiming to be active.[4] Finally, a pure index fund has almost zero tracking error and Active Share.

This methodology is used to characterize active management for all-equity mutual funds in the US. First, we determine how much and what type of active management each fund practices, and we test how this is related to other fund characteristics such as size, fees, flows, and prior returns. Second, we examine the time series from 1980 to 2003 to understand the evolution of active management over time. Third, we investigate fund performance to find out whether more active managers have more skill and whether that skill survives their fees and expenses. Our methodology allows us to focus on the performance of the truly active

[3] In principle, either dimension could be measured entirely from portfolio holdings or from returns. For example, we also use industry-level Active Share in this paper as a holdings-based proxy for industry bets.

[4] Fidelity Magellan at the end of our sample period is one of the most prominent examples, despite the denials by its manager (e.g. *The Wall Street Journal*, 5/28/2004, "Magellan's Manager Has Regrets").

funds as well as the different types of active funds, complementing the existing mutual fund literature which has largely treated all mutual funds as one homogeneous group.

In the cross section of funds, we find wide dispersion along both dimensions of active management. For example, a tracking error of 4-6% can be associated with an Active Share anywhere between 30% and 100%, thus including both closet indexers as well as very active funds. The Active Share of an individual fund is extremely persistent over time.

Consistent with the popular notion, small funds are indeed more active than large funds; however, the effect is economically small, and it only becomes significant after about $1bn in assets. The expense ratio is much lower for index funds, but for all other funds it exhibits surprisingly little relationship to Active Share, which makes closet indexers disproportionately expensive.

The fraction of pure index funds grew substantially over the 1990s, from about 1% to 15% of mutual fund assets. However, the fraction of closet indexers increased even more significantly: funds with low Active Share (20%-60%) had about 30% of all assets in 2003, compared with almost zero in the 1980s. This trend dragged down the average Active Share of non-index large-cap funds from about 80% to 60% over the same period.

Fund performance is significantly related to active management, as revealed by a two-dimensional sort of non-index funds by Active Share and tracking error. Funds with the highest Active Share exhibit some skill and pick portfolios which outperform their benchmarks by 1.51%-2.40% per year. After fees and transaction costs, this outperformance decreases to 1.13%-1.15% per year. In contrast, funds with the lowest Active Share have poor benchmark-adjusted returns and alphas before expenses (between 0.11% and -0.63%) and do even worse after expenses, underperforming by -1.42% to -1.83% per year. The differences in performance across the top and bottom Active Share groups are also statistically significant.

Interestingly, tracking error by itself is *not* related to fund returns. Hence, not all dimensions of active management are rewarded in the market, but the dimension captured by Active Share is.

Economically, these results suggest that the most active stock pickers have enough skill to outperform their benchmarks even after fees and transaction costs. In contrast, funds focusing on factor bets seem to have zero to negative skill, which leads to particularly bad performance after fees. Hence, it appears that there are some mispricings in individual stocks that active managers can exploit, but broader factor portfolios may either be too

efficiently priced or too difficult for the managers to predict. Closet indexers, unsurprisingly, exhibit zero skill but underperform because of their expenses.

Active Share is very significantly related to performance within the smallest 60% of funds, producing a spread in returns of 2.5%-3.8%. A weaker but still positive relationship exists for the largest 40% of funds, where the return spread varies from 1% to 2% per year.

Among the highest Active Share quintile, there is significant persistence in fund performance even *after* controlling for momentum. The funds in the highest Active Share and highest prior-year return quintiles continue to outperform their benchmarks by 5.10% per year ($t = 3.67$) after expenses, or 3.50% per year ($t = 3.29$) under the four-factor model of Carhart (1997).

The current mutual fund literature has done little to investigate active management per se. Instead, a large volume of research has focused on fund performance directly.[5] For example, a comprehensive study by Wermers (2000) computes mutual fund returns before and after expenses; our results refine those performance results by dividing funds into various active management categories. Even more closely related, Wermers (2003) investigates active management and fund performance but uses only the S&P 500 tracking error as a measure of active management; we add the Active Share dimension, which turns out to be crucial for fund returns, and we use a variety of actual stock market indexes rather than only the S&P 500.

Kacperczyk, Sialm, and Zheng (2005) ask a related question about whether industry concentration of mutual funds explains fund performance. This amounts to testing whether funds with concentrated stock picks or large factor bets in industries perform better than other funds. Our performance results address the broader question about whether any active stock picks are reflected in fees and alphas, and whether any types of factor bets, including ones unrelated to specific industries, are similarly reflected in performance.

Another important feature separating our paper from many others in the literature is the data. First, we have holdings data for the most common benchmark indexes used in the industry over the sample period: the S&P 500, S&P 400, S&P 600, S&P500/Barra Value, S&P500/Barra Growth, Russell 1000, Russell 2000, Russell 3000, Russell Midcap, the value and growth components of the four Russell indexes (i.e., eight Russell style indexes),

[5] Various performance measures have been developed and applied by Jensen (1968), Grinblatt and Titman (1989, 1993), Gruber (1996), Daniel, Grinblatt, Titman, and Wermers (1997), Wermers (2000), Pastor and Stambaugh (2002), Cohen, Coval, and Pastor (2005), and many others. Studies focusing on performance persistence include, for example, Brown and Goetzmann (1995), Carhart (1997), Bollen and Busse (2004), and Mamaysky, Spiegel, and Zhang (2006).

Wilshire 5000, and Wilshire 4500, for a total of 19 indexes. This allows us to compute Active Share relative to a fund's *actual* benchmark index as opposed to picking the same market index for all funds.[6] Second, we use *daily* data on mutual fund returns. This is important for the accurate calculation of tracking error, especially when funds do not keep their styles constant over the years or when funds have only short return histories.

The paper proceeds as follows. Section 2 examines our definition and measures of active management. Section 3 describes the data sources and sample selection criteria. The empirical results for active management are presented in Section 4 and for fund performance in Section 5. Section 6 concludes. All tables and figures are in the appendix.

2 Definition and Measures of Active Management

"Passive management" of a portfolio is easy to define: it consists of replicating the return on an index with a strategy of buying and holding all (or almost all) index stocks in the official index proportions.[7]

"Active management" can then be defined as any deviation from passive management. Measuring it involves measuring the "degree of deviation" from passive management. However, there are different *types* of active management, and this is where the difficulties arise: how to measure the deviation depends on what aspect of active management we want to capture.

2.1 Tracking Error

Tracking error (or more formally, tracking error volatility) is commonly defined[8] as the time-series standard deviation of the difference between a fund return ($R_{fund,t}$) and its benchmark index return ($R_{index,t}$):

$$\text{Tracking error} = Stdev\left[R_{fund,t} - R_{index,t}\right].$$

---

[6] Furthermore, we do not use the (potentially misleading) self-reported benchmarks of funds; instead we assign to each fund that benchmark index whose holdings are closest to the actual holdings of the fund (i.e., the index producing the lowest Active Share).

[7] Without transaction costs, passive managers could just sit on their portfolios and trade only when the benchmark index changes, and if there are fund inflows or outflows, they could simply scale the exact same portfolio up or down. In reality, attempts to minimize transaction costs will lead to small deviations from the index.

[8] See e.g. Grinold and Kahn (1999).

A typical active manager aims for an expected return higher than the benchmark index, but at the same time he wants to have a low tracking error (volatility) to minimize the risk of significantly underperforming the index. Mean-variance analysis in this excess-return framework is a standard tool of active managers (e.g. Roll (1992) or Jorion (2003)).

The common definition of tracking error effectively assumes a beta equal to one with respect to the benchmark index, and thus any deviation from a beta of one will generate tracking error. In this paper we adopt a slightly modified definition of tracking error, obtained by regressing excess fund returns on excess index returns:

$$\begin{aligned} R_{fund,t} - R_{f,t} &= \alpha_{fund} + \beta_{fund}\left(R_{index,t} - R_{f,t}\right) + \varepsilon_{fund,t} \\ \text{Tracking error} &= Stdev\left[\varepsilon_{fund,t}\right]. \end{aligned}$$

Following from this definition, any persistent allocation to cash or to high-beta or low-beta stocks will not contribute to our measure of tracking error.

2.2 Active Share

Our new intuitive and simple way to quantify active management is to compare the holdings of a mutual fund with the holdings of its benchmark index. We label this measure the Active Share of a fund, and we define it as:

$$\text{Active Share} = \frac{1}{2}\sum_{i=1}^{N}\left|w_{fund,i} - w_{index,i}\right|,$$

where $w_{fund,i}$ and $w_{index,i}$ are the portfolio weights of asset i in the fund and in the index, and the sum is taken over the universe of all assets.[9]

Active Share has an intuitive economic interpretation. We can decompose a mutual fund portfolio into a 100% position in the benchmark index, plus a zero-net-investment long-short portfolio. The long-short portfolio represents all the active bets the fund has taken. Active Share measures the size of that long-short portfolio as a fraction of the total portfolio of the fund. We divide the sum of portfolio weight differences by two so that a

[9] We compute the sum across stock positions only, as we apply the measure exclusively to all-equity portfolios. However, in general one should sum up across *all* positions, including cash and bonds, which may also be part of the portfolio (or part of the index).

If a portfolio contains derivatives, Active Share becomes a more complex but still feasible concept. Then we would have to decompose the derivatives into implied positions in the underlying securities (e.g., stock index futures would be expressed as positions in stocks and cash) and compute Active Share across those underlying securities. Because mutual funds tend to have negligible derivative positions, this is not a concern for us.

fund that has zero overlap with its benchmark index gets a 100% Active Share (i.e., we do not count the long side and the short side of the positions separately).

As an illustration, let us consider a fund with a $100 million portfolio benchmarked against the S&P 500. Imagine that the manager starts by investing $100 million in the index, thus having a pure index fund with 500 stocks. Assume the manager only likes half of the stocks, so he eliminates the other half from his portfolio, generating $50 million in cash, and then he invests that $50 million in those stocks he likes. This produces an Active Share of 50% (i.e. 50% overlap with the index). If he invests in only 50 stocks out of 500 (assuming no size bias), his Active Share will be 90% (i.e., 10% overlap with the index). According to this measure, it is equally active to pick 50 stocks out of a relevant investment universe of 500 or 10 stocks out of 100 – in either case you choose to exclude 90% of the candidate stocks from your portfolio.

For a mutual fund that never shorts a stock and never buys on margin, Active Share will always be between zero and 100%. In other words, the short side of the long-short portfolio never exceeds the long index position. In contrast, the Active Share of a hedge fund can significantly exceed 100% due to its leverage and net short positions in individual stocks.

2.3 Combining Active Share with Tracking Error

Why do we need to know the Active Share of a fund if we already know its tracking error? The main limitation of using tracking error alone is that different *types* of active management will contribute to it differently; active management is not a one-dimensional concept and thus it cannot be completely characterized by a one-dimensional measure.

There are two basic ways an active fund manager can hope to outperform his benchmark index: by stock selection or factor timing. Fama (1972) was an early advocate of this return decomposition, which has spawned a large body of research, including for example the performance attribution methodologies of Brinson, Hood, and Beebower (1986) and Daniel, Grinblatt, Titman, and Wermers (1997).

Stock selection means attempting to pick outperforming stocks relative to a benchmark portfolio with similar exposure to systematic risk. This may include controlling for market beta, book-to-market ratio, market capitalization, momentum, or industry. Factor timing, in contrast, involves taking time-varying positions in broader factor portfolios according to

the manager's views of their future returns.[10] Naturally, a manager can also combine both approaches.

While the prior literature has largely focused on ex post returns and performance attribution, we focus on quantifying an active manager's ex ante attempt to engage in stock selection or factor timing. To capture a manager's efforts in the two dimensions, we need two separate measures. We suggest using Active Share and tracking error together to span these two dimensions of active management.

The main conceptual difference between Active Share and tracking error is that tracking error includes the covariance matrix of returns. As a result, tracking error puts significantly more weight on *correlated* active bets – in other words, bets on systematic factors. This makes tracking error a reasonable proxy for factor timing. In contrast, Active Share puts equal weight on all active bets (relative to the index), regardless of whether the risk in such bets is largely diversified away in a portfolio. Thus it serves as a reasonable proxy for stock selection.

Figure 1 illustrates the economics behind the two-dimensional classification of funds. A diversified stock picker may take large stock-specific active positions within industries, producing a high Active Share. If it simultaneously diversifies its active positions across all industries and does not bear any systematic risk relative to the benchmark index, it will have a low tracking error just like closet indexers. Its high Active Share is far from irrelevant: a manager can only outperform the benchmark index by deviating from it, so this is a direct indication of the fund's active efforts to outperform. Conversely, a fund that is exclusively timing broad factor portfolios but not attempting to choose stocks within such portfolios would have high tracking error and low Active Share.[11]

In principle, we could measure either dimension of active management entirely from portfolio holdings or from portfolio returns. Factor timing could be measured either with tracking error, which emphasizes bets on systematic risk, or with Active Share computed over broad factor portfolios (such as the industry-level Active Share in Section 4.1.4,which is closely related to the Industry Concentration Index of Kacperczyk, Sialm, and Zheng (2005)). Stock selection could be measured either with Active Share (or an intra-industry

[10] This is also known as "tactical asset allocation," and the managers may be labeled "market timers" or "sector rotators."

[11] "Low Active Share" here is relative. A factor timer with a 100% weight in one out of ten industry portfolios has an Active Share of 90% even without any stock selection. But in reality the active factor weights are likely to be less aggressive – and regardless of those factor weights, adding stock selection to factor timing will always increase Active Share even further, though it may hardly affect tracking error.

measure of Active Share), or with residual volatility from a multifactor regression of fund return on a number of systematic factor portfolios (intended to capture all exposure to systematic risk).

The choice of tracking error and Active Share as proxies for the two dimensions of active management brings some clear benefits. Tracking error allows us to measure factor timing without assuming anything about how fund managers define factor portfolios at each point in time, whereas a holdings-based approach would require such assumptions. Tracking error is also by far the most commonly used measure of active management in practice. Active Share also does not require any assumptions about the relevant factor portfolios, and it is an extremely simple and intuitive measure with a convenient economic interpretation.

3 Empirical Methodology

3.1 Data on Holdings

In order to compute Active Share, we need data on the portfolio composition of mutual funds as well as their benchmark indexes. All stock holdings, for both funds and benchmark indexes, are matched with the CRSP stock return database.

The stock holdings of mutual funds are from the CDA/Spectrum mutual fund holdings database maintained by Thomson Financial. The database is compiled from mandatory SEC filings as well as voluntary disclosures by mutual funds. Starting in 1980, it reports most mutual fund holdings quarterly. Wermers (1999) describes the database in more detail.

As benchmarks for the funds, we include essentially all indexes used by the funds themselves over the sample period. We have a total of 19 indexes from three index families: S&P/Barra,[12] Russell, and Wilshire.

The S&P/Barra indexes we pick are the S&P 500, S&P500/Barra Growth, S&P500/Barra Value, S&P MidCap 400, and S&P SmallCap 600. The S&P 500 is the most common large-cap benchmark index, consisting of approximately the largest 500 stocks. It is further divided into a growth and value style, with equal market capitalization in each style, and this forms the Barra Growth and Value indexes which together sum up to the S&P 500. The S&P 400 and S&P 600 consist of 400 mid-cap and 600 small-cap stocks, respectively. The index constituent data for the S&P/Barra indexes are directly from Standard & Poor's. We have month-end constituents for the large-cap style indexes starting in 9/1992; the S&P

[12] The Barra indexes ceased to be the official S&P style indexes as of December 16, 2005, but this is irrelevant for our sample.

400 holdings data start in 7/1991 and the S&P 600 start in 12/1994. The S&P 500 data cover the sample since 1/1980.

From the Russell family we have 12 indexes: the Russell 1000, Russell 2000, Russell 3000 and Russell Midcap indexes, plus the value and growth components of each. The Russell 3000 covers the largest 3,000 stocks in the U.S. and the Russell 1000 covers the largest 1,000 stocks. Russell 2000 is the most common small-cap benchmark, consisting of the smallest 2,000 stocks in the Russell 3000. The Russell Midcap index contains the smallest 800 stocks in the Russell 1000. The index constituent data are from Frank Russell Co. and start in 12/1978.

Finally, we include the two most popular Wilshire indexes (now owned by Dow Jones), namely the Wilshire 5000 and Wilshire 4500. The Wilshire 5000 covers essentially the entire U.S. equity market, with about 5,000 stocks in 2004 and peaking at over 7,500 stocks in 1998. The Wilshire 4500 is equal to the Wilshire 5000 minus the 500 stocks in the S&P 500 index, which makes it a mid-cap to small-cap index. The Wilshire index constituent data are from Wilshire Associates and start in 1/1979.

In order to cover all basic investment styles over our full time period and to keep the set of benchmarks as constant as possible, we use all the data we have, even if it includes constituent data backdated to a time before the inception of an index.[13] This has an effect on our results in the 1980s, but it has no effect on our performance results which start in 1/1990.

3.2 Data on Returns

Monthly returns for mutual funds are from the CRSP mutual fund database. These are net returns, i.e. after fees, expenses, and brokerage commissions but before any front-end or back-end loads. Monthly returns for benchmark indexes are from S&P, Russell, and Ibbotson Associates, and all of them include dividends.

Daily returns for mutual funds are from multiple sources. Our main source is Standard and Poor's which maintains a comprehensive database of live mutual funds.[14] We use their "Worths" package which contains daily per-share net asset values (assuming reinvested dividends) starting from 1/1980. Because the S&P data does not contain dead funds, we

[13] This means that we backdated the benchmark index holdings ourselves (Wilshire 4500 before 1983) or inferred intermediate month-end holdings from officially backdated quarter-end holdings (Russell indexes before 1987).

[14] This is also known as the Micropal mutual fund data.

supplement it with two other data sources. The first one is the CRSP mutual fund database which also contains daily returns for live and dead funds starting in 1/2001. The second one is a database used by Goetzmann, Ivkovic, and Rouwenhorst (2001) and obtained from the Wall Street Web. It is free of survivorship bias and contains daily returns (assuming reinvested dividends) from 1/1968 to 1/2001, so we use it to match dead funds earlier in our sample. Whenever available, we use the S&P data because it appears slightly cleaner than the latter two sources.

Daily returns for benchmark indexes are from a few different sources. The S&P 500 (total return) is from CRSP, while the rest of the S&P, Russell, and Wilshire index returns are directly from the index providers.

3.3 Sample Selection

We start by merging the CRSP mutual fund database with the CDA/Spectrum holdings database. The mapping is a combined version of the hand-mapping used in Cohen, Coval, and Pastor (2005) and the algorithmic mapping used in Frazzini (2005), where we manually resolve any conflicting matches.

For funds with multiple share classes in CRSP, we compute the sum of total net assets in each share class to arrive at the total net assets in the fund. For the expense ratio, loads, turnover, and the percentage of stocks in the portfolio we compute the value-weighted average across the share classes. For all other variables such as fund name, we pick the variables from the share class with the highest total net assets.

We want to focus on all-equity funds, so we require each fund to have a Wiesenberg objective code of growth, growth and income, equity income, growth with current income, maximum capital gains, small capitalization growth, or missing.[15] We also require an ICDI fund objective code of aggressive growth, growth and income, income, long-term growth, or missing.[16] Finally, we require that the investment objective code reported by Spectrum is aggressive growth, growth, growth and income, unclassified, or missing. All these criteria most notably exclude any bond funds, balanced and asset allocation funds, international funds, precious metals, and sector funds.[17]

---

[15] CRSP also has a variable which indicates the type of securities mainly held by a fund, but the data for it is so incomplete as to render the variable much less useful.

[16] The Wiesenberg objective code is generally available up to 1991 and missing in the later part of the sample, while the ICDI objective code is generally available starting in 1992 and missing in the earlier part of the sample.

[17] Many studies exclude sector funds because they may appear very active while in reality they simply

We then look at the percentage of stocks in the portfolio as reported by CRSP, compute its time series average for each fund, and select the funds where this average is at least 80% or missing.[18] Because this value is missing for many legitimate all-equity funds, we also separately compute the value of the stock holdings and their share of the total net assets of the fund.[19] We require the time-series average of the computed equity share of each fund to be at least 80%.[20] This confirms the all-equity focus of the remaining funds, in particular the ones with missing data items.

To compute Active Share, the report date of fund holdings has to match the date of index holdings. For virtually all of our sample this is not a problem: our index holdings are month-end but so are the fund holdings. However, we still drop the few non-month-end observations from our sample.[21]

Evans (2004) discusses an incubation bias in fund returns, which we address by eliminating observations before the starting year reported by CRSP as well as the observations with a missing fund name in CRSP.

We require at least 100 trading days of daily return data for each fund in the 6 months immediately preceding its holdings report date. This is necessary for reasonably accurate estimates of tracking error, but it does decrease the number of funds in our sample by 5.4%. Naturally a larger fraction of funds is lost in the 1980s than in the later part of the sample.

Finally, we only include funds with equity holdings greater than $10 million.

After the aforementioned screens, our final sample consists of 2,647 funds in the period 1980-2003. For each year and each fund, the stock holdings are reported for an average of three separate report dates (rdate); the total number of such fund-rdate observations in the sample is 48,354.

---

invest according to their sector focus, perhaps even passively tracking a sector index. In our study we could include them, but this would require data on all the various sector indexes.

[18] Several all-equity funds have zeros for this variable, so we treat all zeros as missing values.

[19] We include only the stock holdings we are able to match to the CRSP stock files. Total net assets is preferably from Spectrum (as of the report date), then from CRSP mutual fund database (month-end value matching the report date); if neither value is available, we drop the observation from the sample.

[20] To reduce the impact of data errors, we first drop the observations where this share is less than 2% or greater than 200%. For example, some fund-rdates have incorrectly scaled the number of shares or the total net assets by a factor of 0.001.

[21] We require that the reported holdings date is within the last 4 calendar days of the month. This eliminates about 0.01% of the sample.

3.4 Selection of Benchmark Index

Determining the benchmark index for a large sample of funds is not a trivial task. Our solution is to estimate proper benchmark assignment from the data for the full time period from 1980 to 2003.[22] We compute the Active Share of a fund with respect to all 19 indexes and assign the index with the *lowest* Active Share as that fund's benchmark. By construction, this index has the greatest amount of overlap with the stock holdings of the fund across the set of 19 indexes.

Besides being intuitive, our methodology has a few distinct advantages. It cannot be completely off – if it assigns an incorrect benchmark, it happens only because the fund's portfolio actually does resemble that index more than any other index.[23] It also requires no return history and can be determined at any point in time as long as we know the portfolio holdings. Thus we can use it to track a fund's style changes over time, or even from one quarter to the next when a fund manager is replaced.[24]

4 Results: Active Management

In this section we present the empirical results for active management. We start with a cross-sectional analysis of fund characteristics for various types of funds, using the two dimensions of Active Share and tracking error. We then proceed to investigate the determinants of Active Share in a more general multivariate case. Finally, we discuss the time-series evolution of active management.

---

[22] Since 1998, the SEC has required each fund to report a benchmark index in its prospectus; however, this information is not part of any publicly available mutual fund database, and prior to 1998 it does not exist for all funds. These self-declared benchmarks might even lead to a bias: some funds could intentionally pick a misleading benchmark to increase their chances of beating the benchmark by a large margin. This is discussed in Sensoy (2006).

Typically mutual funds have just one benchmark index, but in some cases a fund's objective may justify a split benchmark between two indexes. We do not consider that extension in this paper.

[23] Contrast this to an alternative estimation method of regressing fund returns on various index returns and assigning the index with the highest correlation with the fund. Because the regression approach is based on noisy returns, we might by chance pick a benchmark index that has nothing to do with the fund's investment policy. Furthermore, we cannot run the regression for new funds with short return histories, or for funds that change their benchmark index over time.

[24] An interesting alternative for defining a benchmark (or style) is presented by Brown and Goetzmann (1997).

4.1 Two-Dimensional Distribution of Funds

We first compile the distribution of all funds in our sample along the two dimensions of Active Share and tracking error, and then investigate how various fund characteristics are related to this distribution. The most recent year for which we have complete data is 2002, so we start our analysis with a snapshot of the cross-section of all funds that year. Panel A of Table 1 presents the number of funds as bivariate distributions and also as univariate marginal distributions along the Active Share and tracking error dimensions.

The distribution of funds clearly reveals a positive correlation between the two measures of active management. Yet within most categories of Active Share or tracking error, there is still considerable variation in the other measure. For example, a tracking error of 4-6% can be associated with an Active Share anywhere between 30% and 100%; and an Active Share of 70-80% can go with a tracking error ranging from 2% to over 14%. This confirms that distinguishing between the two dimensions of active management is also empirically important if we want to understand how much each fund engages in stock selection and factor timing.[25]

Funds with an Active Share less than 20% consist of pure index funds. When we refer to "closet indexers" throughout this paper, we generally mean non-index funds with relatively low Active Share, sometimes specifically referring to the funds with an Active Share of only 20%-60%.[26]

4.1.1 Are Smaller Funds More Active?

Funds with high Active Share indeed tend to be small while funds with low Active Share tend to be larger. Panel B in Table 1 shows that the median fund size varies from less than $200 million for high Active Share funds to $250 million and above for low Active Share

[25] While the Active Share numbers are based on reported fund holdings at the end of a quarter, it is unlikely that any potential "window dressing" by funds would systematically distort their Active Share. For example, to increase Active Share by 10% at the end of each quarter and to decrease it by the same 10% a few days later would require 80% annual portfolio turnover. A fund with an average portfolio turnover of 80% would therefore double its turnover to 160%, incurring large trading costs in the process, all in an effort merely to increase its Active Share by 10%. This seems rather implausible.

[26] It is very hard to see how an active fund could justify investing in more than half of all stocks, because regardless of the managers' beliefs on individual stocks, he must know that at most half of all stocks can beat the market. Thus a fund with an Active Share less than 50% is always a hybrid between a purely active and purely passive portfolio.

funds. The relationship is almost monotonic when going from the most active funds to closet indexers: fund size is indeed negatively correlated with active management.

Figure 2 shows scatter plots of Active Share as a function of fund size for all non-index funds in 2002, separated into funds with large-cap, mid-cap, and small-cap benchmarks. It also shows the average Active Share and the Active Share of a marginal dollar added to a fund's portfolio, both computed from a non-parametric kernel regression of Active Share on log fund size.[27]

The first interesting finding is that the Active Share of a marginal dollar given to a fund begins to fall only for funds above $1 billion. Second, closet indexers (non-index funds with Active Share below 60%) are almost exclusively benchmarked to a large cap index. The average Active Share for both small-cap and mid-cap funds remains above 80% even for the largest funds.

For example, for funds benchmarked to a large-cap index the Active Share of that marginal dollar stays constant at roughly 70% for all the way from a $10 million fund to a $1 billion fund, meaning that these small-to-medium-sized active large-cap funds tend to index approximately 30% of their assets. Above $1 billion in assets Active Share starts to fall more rapidly, first to 60% at $10 billion and then to about 50% for the largest funds, implying that the largest large-cap funds index about one half of their new assets.

However, we should be somewhat cautious when interpreting these results for an individual fund. There is substantial dispersion in Active Share for all fund sizes, so while the mean is descriptive of the entire population, many individual funds still deviate from it significantly in either direction.

Finally, our calculations for Active Share put us in a unique position by allowing us to test one of the assumptions of a recent and prominent theoretical model by Berk and Green (2004), who predict a strong relationship between fund size and active management. In the model, an active manager typically starts with the ability to generate a positive alpha, but he also faces a linear price impact[28] which reduces his initial alpha. The manager then optimally chooses the size of his active portfolio to maximize his dollar alpha, implying that *all* the remaining assets in the fund will be indexed. In other words, once a fund has reached some minimum size, the active share of a marginal dollar should be *zero*.

Figure 2 shows that marginal Active Share is instead almost equal to the average Active

[27] We use the Nadaraya-Watson kernel estimator with a Gaussian kernel and a bandwidth equal to 0.5. Other reasonable bandwidths give similar results.

[28] This in turn generates a quadratic dollar cost.

Share, about 70% for most large-cap funds. The regression evidence in Section 4.2 further shows that recent inflows of assets do not have any economically meaningful impact on the Active Share of a fund. *Qualitatively* it is still true that Active Share decreases with fund size, but *quantitatively* it is very hard to reconcile this result with the zero marginal Active Share implied by the model.

In fact, Figure 2 suggests an alternative story: when a fund receives inflows, instead of indexing all the new assets, it simply scales up its existing positions. This too is a simplification, but it would match the data on active positions much better. It is also supported by Pollet and Wilson (2006) who find that "funds overwhelmingly respond to asset growth by increasing their [existing] ownership shares rather than by increasing the number of investments in their portfolio."

Nevertheless, this evidence should not be interpreted as a complete rejection of the economics in Berk and Green (2004). These new empirical results simply suggest a reevaluation of the exact mechanism behind their intuitively appealing story.

4.1.2 Fees and Closet Indexing

Panel A of Table 2 shows the equal-weighted expense ratio of all funds across Active Share and tracking error in 2002. The equal-weighted expense ratio across all funds in the sample is 1.24% per year, while the value-weighted expense ratio (unreported) is lower at 0.89%.

Index funds clearly have the lowest expense ratios. The equal-weighted average of the lowest Active Share and tracking error group is 0.47% per year.[29]

The funds with the highest Active Share charge an average expense ratio of 1.42%. The other active fund groups exhibit slightly lower fees for lower Active Shares, but the differences are economically small for these intermediate ranges of Active Share. For example, the average expense ratio for funds with Active Share between 30% and 40% is about 1.08% per year, which is closer to the 1.23% of the group with Active Share between 60% and 70% than the 0.47% of the pure index funds, but also clearly lower than the average expense ratio of 1.42% for the funds with the highest Active Share.

[29] The value-weighted average is only 0.22%, which indicates that especially the largest index funds have low fees.

4.1.3 Portfolio Turnover

Portfolio turnover[30] for the average mutual fund is 95% per year (Table 2, Panel B). Average turnover for fund groups varies from 18% for index funds to 210% for one of the highest Active Share groups.

The table reveals a surprisingly weak positive correlation between Active Share and turnover.[31] Almost all non-index fund groups have roughly comparable turnover averages, while the index funds clearly stand out with their lower turnover. This would be consistent with closet indexers (perhaps unwittingly) masking their passive strategies with portfolio turnover, i.e. a relatively high frequency of trading their rather small active positions. Tracking error turns out to predict turnover better than Active Share, implying that the strategies generating high tracking error also involve more frequent trading.

4.1.4 Industry Concentration and Industry-Level Active Share

So far we have computed Active Share at the level of individual stocks. If we compute Active Share at the level of industry Portfolios, the resulting "industry-level Active Share" indicates the magnitude of active positions in entire industries or sectors of the economy. If we contrast this measure with Active Share, we can see how much each fund takes industry bets relative to its bets on individual stocks.

We assign each stock to one of ten industry portfolios. The industries are defined as in Kacperczyk, Sialm, and Zheng (2005).

Table 3 shows the industry-level Active Share across the Active Share and tracking error groups. Within tracking error group, industry-level Active Share is relatively constant even as stock-level Active Share varies from 50% to 100%, especially for tracking error between 6% and 14%. Within Active Share groups , industry-level Active Share increases significantly with tracking error.

This confirms our earlier conjecture that high tracking error often arises from active bets on industries, whereas active stock selection without industry exposure allows tracking error to remain relatively low.

---

[30] CRSP defines the "turnover ratio" of a fund over the calendar year as the "minimum of aggregate purchases of securities or aggregate sales of securities, divided by the average Total Net Assets of the fund."

[31] Using pooled annual observations, the correlation of turnover with Active Share is 18% (with Spearman rank correlation at 17%). Correlations within individual years are similar.

4.2 Determinants of Active Share

To complement the nonparametric univariate results, we run a panel regression of Active Share on a variety of explanatory variables (Table 4). Since some variables are reported only annually, observations are at the fund-year level; when a fund has multiple holdings report dates during the year, we choose the last one.

As independent variables we use tracking error, turnover, expense ratio, and the number of stocks, which are all under the fund manager's control and thus clearly endogenous, as well as fund size, fund age, manager tenure, prior inflows,[32] prior benchmark returns, and prior benchmark-adjusted returns, which are beyond the manager's direct control. We also include year dummies to capture any fixed effect within the year. Because both Active Share and many of the independent variables are persistent over time, we cluster standard errors by fund.

We find that tracking error is by far the most closely related to Active Share: it explains about 13% of the variance in Active Share (the year dummies explain about 10%). Economically, its coefficient of 1.8 (column 2) means that a 5% increase in annualized tracking error increases Active Share by about 9%. This is significant, but it still leaves a great deal of unexplained variance in Active Share.

Fund size is related to Active Share, although this relationship is nonlinear and economically not strong. The expense ratio is statistically significant, but the effect is also economically small: a (large) 1% increase in expense ratio increases Active Share by only about 4.4%. In a similar fashion, turnover has some statistical but no economic significance. Interestingly, fund age and manager tenure act in opposite directions, where long manager tenure is associated with higher Active Share.

Fund inflows over the prior one to three years do not matter for Active Share. This may appear surprising, but it only means that when managers get inflows, they quickly reach their target Active Share, and thus prior fund flows add no explanatory power beyond current fund size. This result is not affected by the presence of control variables (such as prior returns) in the regression.

Benchmark-adjusted returns over the prior three years are significantly related to Active Share. This suggests that fund managers that were successful in the past choose a higher Active Share.

The return on the benchmark index from year $t-3$ to $t-1$ is related to lower Active Share.

[32] Cumulative percentage inflow over the prior one year, and the preceding two years, winsorized at the 1st and 99th percentiles.

In other words, funds are most active when their benchmark index has gone down in the past few years *relative* to the other indexes. Note that the regression includes year dummies, so the effect is truly cross-sectional and not explained by an overall market reaction.[33]

At a more general level, the regression results reveal that Active Share is not easy to explain with other variables – even the broadest specification produced an R^2 of only 32%. Hence, it is indeed a new dimension of active management which should be measured separately and cannot be conveniently subsumed by other variables.

4.3 Active Management over Time

4.3.1 Active Share

Table 5 shows the time-series evolution of active management from 1980 to 2003, as measured by Active Share. There is a clear time trend toward lower Active Share. For example, the percentage of assets under management with Active Share less than 60% went up from 1.5% in 1980 to 44.8% in 2003. Correspondingly, the percentage of fund assets with Active Share greater than 80% went down from 42.8% in 1980 to 23.3% in 2003.

The fraction of index funds before 1990 tends to be less than 1% of funds and of their total assets but grows rapidly after that. Similarly, there are very few non-index funds with Active Share below 60% until about 1987, but since then we see a rapid increase in such funds throughout the 1990s, reaching about 18% of funds and about 30% of their assets in 2000-2001. This suggests that closet indexing has only been an issue since the 1990s – before that, almost all mutual funds were truly active.

4.3.2 Fund-Level Active Share vs. Aggregate Active Share

Active Share can also be computed for the entire mutual fund sector rather than only for individual funds. This aggregate Active Share indicates whether the entire mutual fund sector can act as a marginal investor, buying underpriced stocks and selling overpriced ones, thus helping to make the cross-section of stock prices more efficient. Furthermore, just like Active Share for individual funds, aggregate Active Share is direct evidence of the potential of the entire mutual fund sector to outperform its benchmarks and add value to its investors.

---

[33] In fact the t-statistics on the benchmark index returns are likely to be somewhat overstated because the benchmark index returns (common to all stocks with the same benchmark) will also capture some benchmark-specific differences in Active Share.

Figure 3 shows the aggregate Active Share for non-index funds, together with the equal-weighted and the value-weighted averages at the individual fund level. To compute aggregate Active Share, we sum up all stock positions across individual funds into one large aggregate fund and then compute the Active Share of that aggregate portfolio. To keep the aggregation meaningful, we do not mix funds with different benchmark indexes; we only use funds benchmarked to the S&P 500 (the most common index) for all three time series.

If funds never take active positions against each other, the value-weighted average Active Share should equal the aggregate Active Share. If instead they trade only against each other, e.g. if these funds were the only investors in the market, the aggregate Active Share should sum up to zero. The figure shows that about one half of those active positions actually cancel out each other: in the 1980s, the aggregate Active Share falls to about 45% from a value-weighted average of 75-80%, while in the most recent years the aggregate value has been about 30% out of a fund-level average of 55-60%.

This means that the mutual fund sector as a whole gives investors an Active Share of no more than 30%. The remaining active bets are just noise between funds which will not contribute to an average alpha; any benefit from such bets for one fund must come at the expense of other funds. This helps us understand why the average mutual fund underperforms net of fees: given their low aggregate Active Share, they would have to display considerable skill in their aggregate active bets to fully overcome their fees and expenses.

However, given the large size of the mutual fund sector, their aggregate active bets are still significant in absolute terms, giving mutual funds the potential to bring prices closer to fundamental values. Their performance seems consistent with this, with most empirical evidence in the literature finding slight outperformance (before expenses) for mutual fund portfolios.[34]

4.3.3 Persistence of Fund-Level Active Share

Active Share tends to be highly persistent. Each year we rank all funds into Active Share deciles. For all the stocks in each decile, we compute the average decile rank one to five years later. The decile ranking does not change much from year to year: the top decile ranking falls from 10 to 9.67 and the bottom decile rises from 1 to 1.27. Even over five

[34] Equilibrium asset pricing implications due to the presence of financial institutions such as mutual funds have been explored in a theoretical model by Petajisto (2005). Our empirical estimate for aggregate Active Share can also be used to calibrate that model and to confirm its parameter selection as reasonable.

years, the top decile rank falls only to 8.88 from 10 while the bottom decile rank rises to 2.08 from 1. A decile transition matrix over one year tells a similar story with the diagonal elements ranging from about 40% to 75%. Hence, Active Share this year is a very good predictor of Active Share next year and thereafter.[35]

Tracking error ranks are also persistent but somewhat less so: five years later the top decile has fallen from 10 to 7.52, and the bottom decile has risen from 1 to 2.24.

5 Results: Fund Performance

This section analyses how active management relates to benchmark-adjusted fund returns. We look at both "net returns," which we define as the investors' returns after all fees and transaction costs, and "gross returns," which we define as the hypothetical returns on the disclosed portfolio holdings.[36] The gross returns help us identify whether any categories of funds have skill in selecting portfolios that outperform their benchmarks, and the net returns help us determine whether any such skill survives the fees and transaction costs of those funds.

Prior studies show that the average fund slightly outperforms the market before expenses and underperforms after expenses. Since outperformance can only arise from active management, we hypothesize that there are cross-sectional differences in fund performance: the more active the fund, the higher its average gross return. However, a priori it is not clear how this performance relationship shows up across the two dimensions of active management (i.e., whether Active Share matters more than tracking error) or whether the relationship is linear. For net returns the relationship is even more ambiguous a priori because we do not know how fees and transaction costs are related to the two dimensions of active management.

We pick 1990-2003 as our sample period. This is motivated by Table 5, which confirms that almost all funds were very active in the 1980s. In contrast, starting around 1990 we begin to see some heterogeneity in the distribution, with a meaningful mass of active (non-index) funds having a modest Active Share of 60% or less. It is this cross-sectional dispersion in active management that we conjecture will show up as dispersion in fund performance.

[35] Results are not tabulated to save space and are available upon request.

[36] The same conventions were followed by e.g. Wermers (2000).

Because pure index funds are conceptually different from active funds, we conduct the entire performance analysis only for active (non-index) funds.

5.1 Fund Performance: Active Share vs. Tracking Error

The sample consists of monthly returns for each fund. A fund is included in the sample in a given month if it has reported its holdings in the previous twelve months. Each month we sort funds first into Active Share quintiles and then further into tracking error quintiles. We compute the equal-weighted benchmark-adjusted return within each of the 25 fund portfolios and then take the time series average of these returns over the entire sample period.[37]

Panel A in Table 6 shows the average benchmark-adjusted net returns on these fund portfolios. When we regress the monthly benchmark-adjusted returns on the four-factor model of Carhart (1997), thus controlling for exposure to the market, size, value, and momentum, we obtain the alphas shown in Panel B.

The average fund loses to its benchmark index by 0.43% per year, and the loss increases to 1.14% when controlling for the four-factor model. Tracking error does not help us much when picking funds: the marginal distribution across all tracking error quintiles shows consistently negative benchmark-adjusted returns and alphas. Going from low to high tracking error may even hurt performance, which is statistically significant for the lowest Active Share groups.

In contrast, Active Share *does* improve fund performance relative to the benchmark. The difference in benchmark-adjusted return between the highest and lowest Active Share quintiles is 2.55% per year ($t = 3.47$), which further increases to 2.98% ($t = 4.51$) with the four-factor model. The difference in abnormal returns is positive and economically significant within all tracking error quintiles. An investor should clearly avoid the lowest three Active Share quintiles and instead pick from the highest Active Share quintile. Funds in the highest Active Share quintile beat their benchmarks by 1.13% ($t = 1.60$), or 1.15% ($t = 1.86$) with the four-factor model.

Panels A and B in Table 7 report the corresponding results for gross returns. The high Active Share funds again outperform the low Active Share funds with both economical and statistical significance. The benchmark-adjusted returns indicate that the lowest Active

[37] Since we have so many portfolios of funds, we do not use value weights. In some years the largest funds each account for about 4% of all fund assets, so a value-weighted portfolio return could end up being essentially the return on just one fund. We later also sort funds explicitly on size.

Share funds essentially match their benchmark returns while the highest Active Share funds beat their benchmarks by 2.40% per year ($t = 2.80$). The four-factor model reduces the performance of all fund portfolios but does not change the difference in returns across Active Share and still leaves an economically significant 1.51% outperformance for the highest Active Share funds ($t = 2.23$). Tracking error again exhibits a zero to negative (but statistically insignificant) relationship to fund performance.

The evidence in these two panels suggests that the funds with low Active Share and high tracking error tend to do worst, both in terms of net and gross returns, which implies that factor bets tend to destroy value for fund investors. Closet indexers (low Active Share, low tracking error) also exhibit no ability and tend to lose money after fees and transaction costs.

The best performers are concentrated stock pickers (high Active Share, high tracking error), followed by diversified stock pickers (high Active Share, low tracking error). Both groups appear to have stock-picking ability, and even after fees and transaction costs the most active of them beat their benchmarks.

If we reverse the order of sorting, the results are similar: Active Share is related to returns even within tracking error quintiles, while tracking error does not have such predictive power. A separate subperiod analysis of 1990-1996 and 1997-2003 produces very similar point estimates for both seven-year periods, so the results seem consistent over the entire sample period.

Our general results about the profitability of stock selection and factor timing agree with Daniel, Grinblatt, Titman, and Wermers (1997), who find that managers can add value with their stock selection but not with their factor timing. Because we develop explicit measures of active management, we can refine their results by distinguishing between funds based on their degree and type of active management, thus establishing the best and worst-performing subsets of funds.

We also complement the work of Kacperczyk, Sialm, and Zheng (2005) who find that mutual funds with concentrated industry bets tend to outperform. Their Industry Concentration Index is highest among the concentrated stock pickers and lowest among the closet indexers, with the diversified stock picks and factor bets in the middle. As our paper adds a second dimension of active management, we can further distinguish between these middle groups of funds. This is important for performance because the diversified stock picks outperform and factor bets underperform; consequently, Active Share turns out to

be the dimension of active management that best predicts performance. We discuss the comparison in more detail in Section 5.6.

Part of the difference in net return between the high and low Active Share funds arises from a difference in the "return gap" of Kacperczyk, Sialm, and Zheng (2006). This accounts for 0.64% of the 2.55% spread in benchmark-adjusted net return and 1.22% of the 2.98% spread in four-factor alphas. Hence, if the high Active Share funds have higher trading costs, this is more than offset by the funds' short-term trading ability and their other unobserved actions. Yet most of the net return difference between the high and low Active Share funds still comes from the long-term performance of their stock holdings.

The four-factor betas for the portfolios when their benchmark-adjusted returns are regressed on the market excess return, SMB, HML, and UMD (momentum portfolio)[38] are small on average (-0.01, 0.11, 0.05, and 0.02, respectively), which means that funds collectively do not exhibit a tilt toward any of the four sources of systematic risk. Across Active Share groups, there is no pattern in any of the betas. Across tracking error groups there is more variation in systematic risk: funds with high tracking error tend to be more exposed to market beta and small stocks, with slight preferences for growth stocks and momentum. This exposure seems natural because systematic risk is precisely what produces a high tracking error for a fund.

5.2 Fund Size and Active Share

Since fund size is related to both active management and fund returns, we next investigate how size interacts with Active Share when predicting fund returns. We sort funds into quintiles first by fund size and then by Active Share. The results are reported in Table 8. The median fund sizes for the size quintiles across the sample period are \$28M, \$77M, \$184M, \$455M, and \$1,600M.[39]

Controlling for size, Active Share again predicts fund performance. Within the smallest fund size quintile, the difference between net benchmark-adjusted returns for the top versus the bottom Active Share quintiles equals 2.92% per year and 3.78% after adjusting for the four-factor model. Even within the next two size quintiles, the difference in net performance varies from 2.53% to 3.20% and maintains its statistical significance. For the second-largest fund quintile the difference is slightly lower, ranging from 1.72% to 1.83% per year, and

[38] Results available upon request.

[39] In the 25 basic portfolios sorted on Active Share and tracking error, median fund size varies from about \$100M to \$400M.

is still statistically significant. For the largest fund quintile, the difference is lower still at about 1.01% per year and is no longer statistically significant.

Therefore, it is especially for the smaller funds (i.e., excluding the largest 40% of funds) that the highest Active Share funds exhibit economically significant stock-picking ability: their stock picks outperform their benchmarks by about 2.5-3.8% per year, net of fees and transaction costs.

Fund size alone is also negatively related to fund performance: the difference between the smallest versus the largest size quintile in benchmark-adjusted net abnormal returns is 1.01% ($t = 3.05$). This is consistent with the findings of Chen, Hong, Huang, and Kubik (2004). However, fund size is helpful mostly in identifying the funds that underperform (the largest funds); even the smallest funds on average still do not create value for their investors. To identify funds that actually outperform, we also need to look at Active Share.

5.3 Active Share and Performance Persistence

If some managers have skill to beat their benchmark, we would expect persistence in their performance. This persistence should be strongest among the most active funds. To investigate this, we sort funds into quintiles first by Active Share and then by each fund's benchmark-adjusted gross return over the prior one year. We report the results in Table 9

The benchmark-adjusted net returns (Panel A) of the most active funds show remarkable persistence: the spread between the prior-year winners and losers is 6.81% per year ($t = 3.35$). In contrast, the least active funds have a spread of only 1.69% per year ($t = 1.91$). Most interestingly, controlling for the four-factor model that includes momentum, the spread between prior-year winners and losers for the most active funds decreases but remains economically and statistically very significant at 4.48% per year ($t = 3.06$). In contrast, the spread between prior-year winners and losers for the least active funds decreases to 0.47% per year and is no longer significant, which is consistent with the results of Carhart (1997).

From an investor's point of view, the prior one-year winners within the highest Active Share quintile seem very attractive, with a benchmark-adjusted 5.10% ($t = 3.67$) annual net return and a 3.50% ($t = 3.29$) annualized alpha with respect to the four-factor model. The performance of this subset of funds is also clearly statistically significant, supporting the existence of persistent managerial skill.

If we run the same analysis only for below-median size funds, the top managers emerge as even more impressive.[40] Their benchmark-adjusted net returns are 6.49% ($t = 4.40$), or

[40] The full table of results available upon request.

4.84% ($t = 4.04$) after controlling for the four-factor model. This suggests that investors should pick active funds based on all three measures: Active Share, fund size, and prior one-year return.

5.4 Benchmark Performance and Active Share

One benefit of Active Share is that it provides a relatively accurate estimate of each fund's official benchmark index. This allows us to directly compare the performance of each fund to that of its benchmark index; after all, that benchmark through a low-cost index fund is the most direct investment alternative for a mutual fund investor.

The benchmark adjustment is particularly useful as some general investment styles or benchmarks generate nonzero alphas under the four-factor model. For example, the S&P 500 and Russell 2000 indexes both have economically and statistically significant annual four-factor alphas of 1.08% ($t = 2.72$) and -2.73% ($t = -2.58$), respectively, over our performance sample period of 1990-2003.[41] Yet it would be inappropriate to attribute the 1.08% alpha to the "skill" of a purely mechanical S&P 500 index fund; instead, it suggests a misspecification in the four-factor model, which in turn can be addressed by the benchmark-adjustment.

Table 10 shows the performance of the benchmark indexes themselves for the funds in the double sort on Active Share and tracking error. Excess returns relative to the risk-free rate (Panel A) are very similar across Active Share quintiles, with none of the differences statistically significant. This indicates that benchmark-adjustment alone does not affect performance across Active Share quintiles over our sample period. However, the four-factor alphas of benchmark excess returns (Panel B) vary widely: the benchmarks of funds in the highest Active Share quintile have annualized alphas that are 2.75% ($t = 2.53$) lower than the benchmark alphas in the lowest Active Share quintile. This is mostly driven by small-cap funds choosing a higher Active Share and large-cap funds choosing a lower Active Share.

As the difference in benchmark alphas is economically large, it is important to verify that Active Share remains related to abnormal fund returns after controlling for the particular benchmark used. We explore this in the next section.

[41] These are the most common benchmark indexes, but other benchmarks also have large nonzero four-factor alphas over the sample period, varying from about -4% to $+3\%$ per year.

5.5 Fund Performance in a Multivariate Regression

To better isolate the effect of different fund characteristics on fund performance, we run pooled panel regressions of fund performance on all the explanatory variables (Table 11). The values for the independent variables are chosen at the end of each year, while the dependent variable is performance over the following year. We use three different performance metrics: four-factor alphas of benchmark-adjusted returns, four-factor alphas of excess returns (relative to the risk-free rate), and the Characteristic Selectivity (CS) measure of Daniel, Grinblatt, Titman, and Wermers (1997). The first two metrics are calculated using net returns whereas the CS measure is based on gross returns.[42]

The list of explanatory variables includes Active Share, tracking error, turnover, expense ratio, the number of stocks, fund size, fund age, manager tenure, prior inflows, prior benchmark returns, and prior benchmark-adjusted returns. As Active Share was shown earlier to be more strongly related to fund performance for smaller funds, we also include the interaction of Active Share with a dummy variable indicating below-median fund size for that year. Since the pooled panel regressions exhibit significant residual correlations within a year, we also include year dummies and cluster standard errors by year in each regression. [43]

Active Share comes up as a highly significant predictor of future benchmark-adjusted net alphas (column 1), with a coefficient of 0.0722 and a t-statistic of 2.42. This means that, controlling for the other variables, a 30% increase in Active Share is associated with an increase of 2.17% in benchmark-adjusted alpha over the following year. Rather than being subsumed by other variables, the predictive power of Active Share actually goes up when those other variables are added.

Unlike Active Share, tracking error produces a small negative effect on future performance, which is marginally statistically significant.

Size and expenses emerge as the most significant other predictors of returns. Size enters in a nonlinear but economically and statistically significant way, showing that larger funds in our sample tend to underperform.[44,45]

---

[42] Using gross returns instead of net returns for columns 1-4 yields very similar results.

[43] Clustering standard errors by year results in much more conservative standard errors than clustering by fund.

[44] The effects of firm size are not very robust. For example, if the interaction of Active Share and a below-median fund size dummy is added, fund size by itself loses its significance. The same thing happens when the dependent variable is the four-factor alpha (columns 3 and 4).

[45] Prior one-year benchmark-adjusted net returns predict higher future returns, with a 10% outperfor-

Column 2 confirms that the relation between Active Share and performance is stronger for smaller funds, as the interaction of Active Share and a dummy for below-median fund size has a positive and significant coefficient. Overall, for below-median sized funds, a 30% increase in Active Share is associated with an increase in benchmark-adjusted alphas of 2.41% per year.

In columns 3 and 4, our performance measure is the four-factor alpha over net excess fund returns over the risk-free rate, thus without the benchmark-adjustment. As discussed in the previous section, it is important to control both for non-zero alphas of the benchmarks themselves and for differences in benchmarks across Active Share levels. We aim to account for both by adding dummy variables for all 19 benchmarks. The main result is that the predictive ability of Active Share decreases but remains economically and statistically significant at the 10% level.

As a final robustness check, we use a fund's Characteristic Selectivity (CS) measure of performance in columns 5 and 6.[46] For below-median size funds, there is a positive and significant coefficient, indicating that a 30% increase in Active Share is associated with an increase in next year's CS of 0.87%. For above-median size funds, the predictive ability is about half that but no longer statistically significant.

5.6 Comparison of All Measures of Active Management

Table 12 shows a comparison of Active Share with other measures of active management. We compare Active Share with tracking error, industry-level Active Share, Industry Concentration Index, stock concentration index, and turnover.[47] Industry-level Active Share is computed similarly to Active Share, except that it replaces individual stocks with 10 industry portfolios (as in Section 4.1.4). The Industry Concentration Index is computed as in Kacperczyk, Sialm, and Zheng (2005) (see also Section A.1 of this paper), except that the benchmark index is selected following the methodology of our paper. The stock

---

mance producing a 0.75% outperformance the following year, though benchmark-adjusted returns two and three years ago predict returns negatively today. However, neither of these effects is statistically significant. The number of stocks is positively related to future fund performance, such that an additional 100 stocks in the portfolio are associated with an increase in performance of 0.43%. Finally, fund age is a slightly negative predictor of returns, but only if the manager is new. Controlling for other characteristics, a fund that ages together with the manager does not suffer from lower returns.

[46] We again add benchmark dummies to control for non-zero CS of the benchmarks and for differences in benchmarks across Active Share levels.

[47] The Pearson correlations of Active Share with these other measures are 45%, 62%, 36%, 49%, and 18%, respectively, and the Spearman rank correlations are 56%, 60%, 60%, 60%, and 17%.

concentration index is analogous to the Industry Concentration Index, except that it uses individual stocks rather than industry portfolios. Standard errors are clustered by year to be conservative, and all regressions include year dummies.

We first consider each measure's ability to predict performance measured as next year's benchmark-adjusted four-factor alpha from net returns (columns 1-8). In the univariate regressions without control variables, Active Share, industry-level Active Share, and Industry Concentration Index all come up as significant, while tracking error, stock concentration index, and turnover are not significant. When all the variables are included in the same regression (column 7), Active Share dominates the other variables, especially after we add the control variables of Table 11 (column 8). In fact, Active Share is the only variable that is highly significant and remains so in all the regression specifications. Tracking error has a negative sign that is marginally significant, and the (marginal) predictive ability of the Industry Concentration Index disappears after adding the controls.

Finally, columns 9 and 10 use four-factor alphas from excess net fund returns over the risk-free rate, again also adding benchmark dummies. Active Share now predicts performance only for below-median size funds.[48]

6 Conclusions

Traditionally the degree of active management is quantified along just one dimension: tracking error relative to a benchmark index. Yet this fails to capture the two different dimensions of active management: stock selection and factor timing.

This paper points out that active management should be measured in two dimensions with Active Share and tracking error as convenient empirical proxies. Tracking error measures the volatility of portfolio return around a benchmark index, thus emphasizing *correlated* active bets such as exposure to systematic factor risk. Active Share measures the deviation of portfolio holdings from the holdings of the benchmark index, placing equal weight on *all* active bets regardless of diversification and thus emphasizing stock selection. This new methodology also allows us to empirically identify different types of active management: diversified stock picks, concentrated stock picks, factor bets, closet indexing, and pure indexing.

Applying this methodology to all-equity mutual funds, we find significant dispersion

[48] The positive and significant coefficient for Industry-level Active Share in column 10 should be interpreted with caution. Industry-level Active Share has a high (85%) correlation with the Industry Concentration Index, so this may be driven by multicollinearity.

along both dimensions of active management. We also confirm the popular belief that small funds are more active, while a significant fraction of large funds are closet indexers. However, this pattern emerges only gradually after $1bn in assets – before that, fund size does not matter much for the fraction of active positions in the portfolio.

There has been a significant shift from active to passive management over the 1990s. Part of this is due to index funds, but an even larger part is due to closet indexers and a general tendency of funds to mimic the holdings of benchmark indexes more closely. Furthermore, about half of all active positions at the fund level cancel out *within* the mutual fund sector, thus making the aggregate mutual fund positions even less active.

Active management, as measured by Active Share, significantly predicts fund performance. Funds with the highest Active Share outperform their benchmarks both before and after expenses, while funds with the lowest Active Share underperform after expenses. In contrast, active management as measured by tracking error does *not* predict higher returns.

The relationship between Active Share and fund returns exists for all fund sizes but it is stronger within the bottom three fund size quintiles than within the top two quintiles. We also find strong evidence for performance persistence for the funds with the highest Active Share, even after controlling for momentum. From an investor's point of view, funds with the highest Active Share, smallest assets, and best one-year performance seem very attractive, outperforming their benchmarks by 6.5% per year net of fees and expenses.

The general reason why we can find strong performance results seems clear: our methodology allows us to distinguish between different *types* of active funds as well as to focus on the ones that are *truly active*. Most existing literature has treated all non-index funds as one homogeneous group, so our methodology could help researchers refine and potentially improve their existing results. Furthermore, our approach will allow researchers to investigate the risk-taking and incentives of mutual fund managers from a new and economically meaningful perspective.

References

[1] Asness, C., 2004, "An Alternative Future," *Journal of Portfolio Management*, 30th Anniversary Issue, 94-103.

[2] Berk, J.B. and R.C. Green, 2004, "Mutual Fund Flows and Performance in Rational Markets," *Journal of Political Economy*, vol. 112, no. 6, 1269-1295.

[3] Bollen, N.P.B. and J.A. Busse, 2004, "Short-Term Persistence in Mutual Fund Performance," *Review of Financial Studies*, vol. 18, no. 2, 569-597.

[4] Brinson, G.P., L.R. Hood, and G.L. Beebower, 1986, "Determinants of Portfolio Performance," *Financial Analysts Journal*, vol. 42, no. 4, 39-44.

[5] Brown, S.J. and W.N. Goetzmann, 1995, "Performance Persistence," *Journal of Finance*, vol. 50, no. 2, 679-698.

[6] Brown, S.J. and W.N. Goetzmann, 1997, "Mutual Fund Styles," *Journal of Financial Economics*, vol. 43, 373-399.

[7] Carhart, M., 1997, "On Persistence in Mutual Fund Returns," *Journal of Finance*, vol. 52, no. 1, 57-82.

[8] Chen, J., H. Hong, M. Huang, and J.D. Kubik, 2004, "Does Fund Size Erode Performance? Organizational Diseconomies and Active Money Management," *American Economic Review*, vol. 94, no. 5, 1276-1302.

[9] Cohen, R.B., J.D. Coval, and L. Pastor, 2005, "Judging Fund Managers by the Company They Keep," *Journal of Finance*, vol. 60, no. 3, 1057-1096.

[10] Daniel, K., M. Grinblatt, S. Titman, and R. Wermers, 1997, "Measuring Mutual Fund Performance with Characteristic-Based Benchmarks," *Journal of Finance*, vol. 52, no. 3, 1035-1058.

[11] Evans, R.B., 2004, "A Does Alpha Really Matter? Evidence from Mutual Fund Incubation, Termination, and Manager Change," working paper.

[12] Fama, E.F., 1972, "Components of Investment Performance," *Journal of Finance*, vol. 27, no. 3, 551-567.

[13] Frazzini, A., 2005, "The Disposition Effect and Underreaction to News," *Journal of Finance*, forthcoming.

[14] Goetzmann, W.N., Z. Ivkovic, and K.G. Rouwenhorst, 2001, "Day Trading International Mutual Funds: Evidence and Policy Solutions," *Journal of Financial and Quantitative Analysis*, vol. 36, no. 3, 287-309.

[15] Grinblatt, M. and S. Titman, 1989, "Mutual Fund Performance: An Analysis of Quarterly Portfolio Holdings," *Journal of Business*, vol. 62, no. 3, 393-416.

[16] Grinblatt, M. and S. Titman, 1993, "Performance Measurement without Benchmarks: An Examination of Mutual Fund Returns," *Journal of Business*, vol. 66, no. 1, 47-68.

[17] Grinold, R.C. and R.N. Kahn, 1999, *Active Portfolio Management*, 2nd Edition, McGraw-Hill.

[18] Gruber, M.J., 1996, "Another Puzzle: The Growth in Actively Managed Mutual Funds," *Journal of Finance*, vol. 51, no. 3, 783-810.

[19] Jensen, M.C., 1968, "The Performance of Mutual Funds in the Period 1945-1964," *Journal of Finance*, vol. 23, no. 2, 389-416.

[20] Jorion, P., 2003, "Portfolio Optimization with Tracking-Error Constraints," *Financial Analysts Journal*, vol. 59, no. 5, 70-82.

[21] Kacperczyk, M.T., C. Sialm, and L. Zheng, 2005, "On Industry Concentration of Actively Managed Equity Mutual Funds," *Journal of Finance*, vol. 60, no. 4, 1983-2011.

[22] Kacperczyk, M.T., C. Sialm, and L. Zheng, 2006, "Unobserved Actions of Mutual Funds," *Review of Financial Studies*, forthcoming.

[23] Mamaysky, H., M. Spiegel, and H. Zhang, 2006, "Improved Forecasting of Mutual Fund Alphas and Betas," Yale ICF working paper.

[24] Pastor, L. and R.F. Stambaugh, 2002, "Mutual Fund Performance and Seemingly Unrelated Assets," *Journal of Financial Economics*, vol. 63, 315–349.

[25] Petajisto, A., 2005, "Why Do Demand Curves for Stocks Slope Down?" Yale ICF working paper.

[26] Pollet, J.M. and M. Wilson, 2006, "How Does Size Affect Mutual Fund Behavior?" working paper.

[27] Roll, R., 1992, "A Mean/Variance Analysis of Tracking Error," *Journal of Portfolio Management*, vol. 18, no. 4, 13-22.

[28] Sensoy, B.A., 2006, "Incentives and Mutual Fund Benchmarks," working paper.

[29] Wermers, R., 1999, "Mutual Fund Herding and the Impact on Stock Prices," *Journal of Finance*, vol. 54, no. 2, 581-622.

[30] Wermers, R., 2000, "Mutual Fund Performance: An Empirical Decomposition into Stock-Picking Talent, Style, Transactions Costs, and Expenses," *Journal of Finance*, vol. 55, no. 4, 1655-1695.

[31] Wermers, R., 2003, "Are Mutual Fund Shareholders Compensated for Active Management 'Bets'?," working paper.

Appendix A: Other Measures of Active Management

A.1 Industry Concentration Index

Kacperczyk, Sialm, and Zheng (2005) investigate a related question about the industry concentration of mutual funds. They call their measure the Industry Concentration Index, which they define as

$$\text{Industry Concentration Index} = \sum_{i=1}^{I} \left(w_{fund,i} - w_{index,i}\right)^2,$$

where $w_{fund,i}$ and $w_{index,i}$ are the weights of industry i in the fund and in the index, and they sum up across I industry portfolios (instead of N individual stocks). They also use the CRSP value-weighted index as their only benchmark. A more fundamental difference between Active Share and the Industry Concentration Index arises from the fact that the latter uses squared weights. For our study, we prefer to use Active Share for three reasons.

First, Active Share has a convenient economic interpretation: it immediately tells us the percentage of a fund that is different from the benchmark index. If the weights are squared, the numerical value loses this interpretation, and its main purpose is then just to rank funds relative to each other.

Second, different funds have different benchmark indexes, yet Active Share can still be easily applied when comparing any two funds: a 90% Active Share means essentially the same thing whether the benchmark is the S&P 500 (with 500 stocks) or the Russell 2000 (with 2,000 stocks). If we square the weights, we lose the ability to make such easy comparisons across indexes because the number of stocks begins to matter. For example, if a fund with the Russell 2000 as a benchmark is likely to have more stocks in its portfolio than a fund with the S&P 500 as a benchmark because the Russell 2000 investment universe contains four times as many stocks, then the typical active weight in a stock will be smaller and thus the sum of squares will be smaller.[49]

Third, the squared weights make the Industry Concentration Index something of a

[49] As an illustration, assume that the Russell 2000 and S&P 500 are equal-weighted indexes. Assume we have an S&P 500 fund which leaves out one half of the index stocks (the "bottom half") and doubles the its portfolio weight on the other half. Assume we have a Russell 2000 fund that does the same thing with its benchmark. Both of these funds have an active share of 50%, but the concentration index is $\sum_{i=1}^{500}\left(\frac{1}{500}\right)^2 = \frac{1}{500}$ for the former and $\sum_{i=1}^{2000}\left(\frac{1}{2000}\right)^2 = \frac{1}{2000}$ for the latter. Hence, scaling the number of stocks in the benchmark by a factor of 4 also scaled up the concentration index by a factor of 4.

hybrid between Active Share and tracking error.[50] However, to get a more complete picture of active management, we need to quantify it along *two* separate dimensions. We therefore pick two measures which are as different from each other as possible, and here Active Share and tracking error seem to satisfy that objective.

A.2 Turnover

Portfolio turnover has also been suggested as a measure of active management. For our purposes it has some significant shortcomings and plays only a minor role in our tests.

Although portfolio turnover implies an action (i.e., trading) by the fund manager, turnover per se cannot add value to a portfolio – only holding a position does. Turnover just measures the frequency of revisions in the manager's active bets (i.e., positions), but it does not measure the activeness of the bets themselves. These are two different kinds of activeness: either "being busy" with the portfolio, or holding positions that differ significantly from the benchmark and thus have a chance to outperform or underperform. This paper focuses on the latter definition.[51]

Fund inflows and outflows can also generate additional turnover which does not tell us anything about the active management of the fund. Furthermore, if turnover is widely used as a measure of active management, less active funds may have an incentive to generate unnecessary trades to appear more active.

Appendix B: Tables

[50] Assume a fund has no systematic risk except for an index beta of 1. Its tracking error is then given by

$$\sigma\left(R_{fund}-R_{index}\right)=\sigma\left(\sum_{i=1}^{N}\left(w_{fund,i}-w_{index,i}\right)R_i\right)=\sqrt{\sum_{i=1}^{N}\left(w_{fund,i}-w_{index,i}\right)^2\sigma_{\varepsilon_i}^2}.$$

If the stocks (or industry portfolios) have a similar idiosyncratic volatility $\sigma_{\varepsilon_i}^2$, then tracking error will be approximately proportional to the square root of the Industry Concentration Index.

[51] To illustrate this, let us consider the famous Legg Mason Value Trust which beat the S&P 500 index 15 years in a row: in 2003 it had an Active Share of 86%, holding only 30 stocks in the portfolio, yet it had a turnover of only 25%. The same year, iShares Russell 2000 index fund had a turnover of 30% because of turnover in the underlying index. The low turnover of Legg Mason Value Trust simply indicates the long investment horizon of its stock picks rather than any adherence to a benchmark index.

Table 1: **All-equity mutual funds in the US in 2002**, sorted by the two dimensions of active management. Active share is defined as the percentage of a fund's portfolio holdings that differ from the fund's benchmark index. It is computed based on Spectrum mutual fund holdings data and index composition data for 19 common benchmark indexes from S&P, Russell, and Wilshire. Tracking error is defined as the annualized standard deviation of the error term when the excess return on a fund is regressed on the excess return on its benchmark index. It is computed based on daily fund returns and daily index returns over a six-month period before the corresponding portfolio holdings are reported. To include only all-equity funds, every fund classified by CRSP as balanced or asset allocation has been removed from the sample. Also sector funds have been eliminated. In Panel A, if a cell has less than 5 observations (fund-dates), it is shown as empty. In Panel B, a statistic must be based on at least 5 funds to be reported.

Panel A: Number of mutual funds

	Tracking error (% per year)								
Active share (%)	0-2	2-4	4-6	6-8	8-10	10-12	12-14	>14	All
90-100			66	125	77	41	22	26	358
80-90		17	100	120	54	24	10	10	336
70-80		26	124	83	27	5	7	10	281
60-70		75	115	41	12		1		247
50-60	3	102	55	15	3				179
40-50	9	66	20						98
30-40	15	27	3						47
20-30	11	4							14
10-20	8								10
0-10	104	4							109
All	150	323	482	388	174	73	41	48	1678

Panel B: Median net asset value ($M)

	Tracking error (% per year)								
Active share (%)	0-2	2-4	4-6	6-8	8-10	10-12	12-14	>14	All
90-100			174	127	97	95	93	51	115
80-90		264	263	177	109	128	206	43	183
70-80		264	163	149	185	237	96	26	156
60-70		379	256	208	180				259
50-60		262	256	257					251
40-50	151	304	281						275
30-40	535	269							267
20-30	200								196
10-20	58								58
0-10	480								480
All	395	303	220	162	110	110	92	48	190

Table 2: **Expense ratios and annual portfolio turnover for all-equity mutual funds in 2002**, sorted by the two dimensions of active management. The measures of active management are computed as before. Turnover is defined by CRSP as the maximum of annual stock purchases and annual stock sales, divided by the fund's total net assets. To include only all-equity funds, every fund classified by CRSP as balanced or asset allocation has been removed from the sample. Also sector funds have been eliminated. To be reported in the table, a statistic must be based on at least 5 funds.

Panel A: Equal-weighted total expense ratio (%)

	Tracking error (% per year)								
Active share (%)	0-2	2-4	4-6	6-8	8-10	10-12	12-14	>14	All
90-100			1.33	1.37	1.51	1.47	1.49	1.50	1.42
80-90		1.30	1.30	1.43	1.44	1.43	1.37	2.11	1.41
70-80		1.19	1.29	1.37	1.33	1.40	1.85	1.34	1.33
60-70		1.10	1.24	1.35	1.37				1.23
50-60		1.04	1.21	1.43					1.14
40-50	1.12	1.08	1.07						1.08
30-40	1.03	1.06							1.08
20-30	0.92								0.88
10-20	0.71								0.75
0-10	0.47								0.47
All	0.62	1.08	1.27	1.39	1.44	1.45	1.54	1.59	1.24

Panel B: Equal-weighted turnover (%)

	Tracking error (% per year)								
Active share (%)	0-2	2-4	4-6	6-8	8-10	10-12	12-14	>14	All
90-100			71.2	101.7	107.8	118.2	140.0	198.5	108.8
80-90		93.5	101.9	133.5	124.5	134.1	210.2	147.5	123.5
70-80		69.3	91.7	98.6	133.8	80.7	74.2	123.9	96.1
60-70		69.0	93.9	107.5	108.0				89.2
50-60		65.5	92.0	87.4					76.8
40-50	57.1	69.7	61.6						67.3
30-40	72.9	117.4							97.8
20-30	141.7								148.9
10-20	60.0								66.1
0-10	18.1								18.4
All	38.2	73.9	89.9	111.1	116.5	119.1	145.3	170.0	94.8

Table 3: **Industry-level Active Share for all-equity mutual funds in 2002**, sorted by the two dimensions of active management. The measures of active management are computed as before. All stocks are assigned into 10 industry portfolios derived from the 49 Fama-French industry portfolios, and industry-level Active Share is computed using these portfolios (instead of individual stocks) as assets. To include only all-equity funds, every fund classified by CRSP as balanced or asset allocation has been removed from the sample. Also sector funds have been eliminated. To be reported in the table, a statistic must be based on at least 5 funds.

	Tracking error (% per year)								
Active share (%)	0-2	2-4	4-6	6-8	8-10	10-12	12-14	>14	All
90-100			21.1	22.7	25.2	30.6	32.9	40.0	25.8
80-90		17.7	19.2	22.4	27.3	26.2	32.1	33.2	22.9
70-80		15.2	17.8	21.0	23.3	27.1	30.3	33.8	20.0
60-70		12.7	16.6	20.7	22.5				16.5
50-60		10.8	15.2	18.9					13.0
40-50	6.0	8.6	15.1						9.9
30-40	4.6	9.0							7.9
20-30	4.2								4.5
10-20	3.0								3.2
0-10	0.6								0.6
All	1.8	11.2	17.8	21.8	25.2	28.6	32.0	36.5	18.1

Table 4: **Determinants of Active Share for all-equity mutual funds in 1992-2003.** The dependent variable is Active Share for each fund-year observation. All the variables are computed as before. Turnover and expense ratio are annualized values. Fund age and fund manager tenure are measured in years. Fund inflows and returns are all cumulative percentages. Index return represents the benchmark assigned to each fund, and return over the index represents a fund's net return (after all expenses) in excess of its benchmark index. Index funds are excluded from the sample. Since the expense ratio and manager tenure are missing before 1992, we limit all specifications to the same time period. Year fixed-effects are included in all specifications. The t-statistics (in parentheses) are based on standard errors clustered by fund.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Tracking error	1.4015	1.8111	1.7002	1.5965	1.5210	1.4439	
	(19.16)	(18.15)	(17.40)	(16.09)	(12.81)	(12.17)	
Turnover				-0.0016		-0.0021	
				(0.65)		(0.66)	
Expenses				4.4359	4.6230	4.6267	7.7859
				(6.33)	(5.28)	(5.33)	(9.72)
$lg(TNA)$			0.0554	0.0601	0.0451	0.0614	0.0389
			(2.96)	(3.16)	(2.02)	(2.87)	(1.62)
$(lg(TNA))^2$			-0.0177	-0.0171	-0.0150	-0.0177	-0.0166
			(4.85)	(4.58)	(3.56)	(4.36)	(3.65)
Number of stocks						-0.0001	
						(2.04)	
Fund age						-0.0005	-0.0003
						(2.26)	(1.06)
Manager tenure						0.0036	0.0041
						(6.72)	(7.00)
Inflow, t-1 to t						0.0052	0.0045
						(1.30)	(1.04)
Inflow, t-3 to t-1						0.0010	0.0019
						(0.94)	(1.53)
Return over index, t-1 to t					0.1068	0.0996	0.1189
					(8.12)	(7.45)	(8.21)
Return over index, t-3 to t-1					0.1103	0.1089	0.1478
					(9.39)	(9.17)	(13.00)
Index return, t-1 to t						0.0655	0.0756
						(5.28)	(6.02)
Index return, t-3 to t-1					-0.0619	-0.0570	-0.0469
					(7.87)	(6.93)	(5.19)
Year dummies	No	Yes	Yes	Yes	Yes	Yes	Yes
N	11,726	11,726	11,726	11,554	8,417	8,320	8,374
R^2	0.1316	0.2373	0.2642	0.2781	0.2984	0.3235	0.2037

Table 5: **All-equity mutual funds in the US in 1980-2003**, sorted by Active Share. Active Share is computed as before. To include only all-equity funds, every fund classified by CRSP as balanced or asset allocation has been removed from the sample. Also sector funds have been eliminated.

Panel A: Percentage of all-equity mutual funds for given level of Active Share

	Active share (%)					Total number
Year	0-20	20-40	40-60	60-80	80-100	of funds
2003	5.9	2.9	15.1	29.2	46.8	2,026
2002	7.1	3.6	16.5	31.5	41.3	1,678
2001	7.0	3.6	16.9	30.9	41.6	1,656
2000	5.8	5.1	17.2	30.4	41.5	1,501
1999	4.6	2.5	15.0	30.3	47.6	1,402
1998	4.7	1.4	9.7	28.9	55.3	1,286
1997	4.5	1.0	6.9	28.9	58.7	1,171
1996	4.2	0.9	5.9	30.1	59.0	974
1995	4.5	1.0	4.8	31.6	58.1	871
1994	5.2	0.8	3.3	27.3	63.5	766
1993	5.0	0.7	3.1	24.7	66.5	650
1992	3.2	0.7	3.9	27.8	64.5	500
1991	2.3	0.9	4.6	33.3	58.9	428
1990	1.2	1.5	4.0	31.7	61.7	340
1989	0.9	1.2	2.4	22.4	73.2	348
1988	1.0	1.2	2.2	20.6	75.0	300
1987	0.7	0.4	1.8	21.4	75.8	285
1986	0.4		0.4	17.2	82.0	252
1985	0.5		1.0	18.0	80.5	219
1984	0.6		1.7	21.7	76.1	181
1983	0.6		1.8	22.0	75.5	163
1982	0.7		2.6	25.2	71.4	139
1981	0.8		2.3	24.2	72.7	132
1980	0.8		1.3	25.5	72.4	126

Panel B: Percentage of all-equity mutual fund assets for given level of Active Share

	Active share (%)					Total assets
Year	0-20	20-40	40-60	60-80	80-100	($bn)
2003	15.3	7.0	22.5	31.9	23.3	1,954
2002	14.7	8.6	26.4	29.3	21.0	1,728
2001	14.5	10.7	22.9	33.8	18.1	1,973
2000	13.3	10.6	22.3	37.2	16.6	2,335
1999	13.6	7.3	27.0	34.7	17.4	1,996
1998	11.6	3.8	24.4	34.2	26.1	1,505
1997	9.3	0.7	17.3	42.9	29.9	1,121
1996	7.8	0.7	12.6	45.3	33.6	774
1995	5.7	0.6	5.8	54.0	33.8	540
1994	5.5	0.7	5.9	45.0	42.9	370
1993	5.4	0.4	6.0	42.2	45.9	311
1992	4.7	0.9	7.5	47.1	39.8	216
1991	3.6	0.9	6.2	51.7	37.5	166
1990	2.3	1.3	9.3	49.0	38.2	119
1989	1.3	1.2	11.0	39.6	46.9	119
1988	1.1	1.1	9.6	39.5	48.8	96
1987	0.9	0.2	7.4	41.8	49.8	108
1986	0.6		0.1	39.7	59.6	82
1985	0.6		0.5	40.8	58.1	61
1984	0.5		1.0	44.9	53.6	47
1983	0.4		1.0	42.6	56.0	44
1982	0.4		2.4	53.7	43.5	26
1981	0.4		2.4	53.4	43.8	27
1980	0.4		1.1	55.7	42.8	25

Table 6: **Net equal-weighted alphas for all-equity mutual funds in 1990-2003**, sorted by the two dimensions of active management. The measures of active management are computed as before. Net fund returns are the returns to a fund investor after fees and transaction costs. Index funds are excluded from the sample. The table shows annualized returns, followed by t-statistics (in parentheses) based on White's standard errors.

Panel A: Benchmark-adjusted return							
Active Share	Tracking error quintile						
quintile	Low	2	3	4	High	All	High-Low
High	0.09	0.39	1.34	2.76	1.05	1.13	0.97
	(0.09)	(0.41)	(1.52)	(2.86)	(0.62)	(1.60)	(0.44)
4	-0.43	-0.15	0.56	0.50	0.76	0.25	1.20
	(-0.61)	(-0.19)	(0.64)	(0.42)	(0.36)	(0.31)	(0.48)
3	-1.42	-0.98	-0.25	-0.49	-0.60	-0.75	0.82
	(-2.06)	(-1.34)	(-0.29)	(-0.45)	(-0.35)	(-0.95)	(0.43)
2	-1.89	-1.14	-1.13	-1.01	-1.66	-1.37	0.23
	(-3.20)	(-1.55)	(-1.53)	(-1.08)	(-1.22)	(-1.99)	(0.16)
Low	-1.35	-1.32	-1.28	-1.51	-1.63	-1.42	-0.28
	(-4.95)	(-3.68)	(-2.77)	(-2.76)	(-2.13)	(-3.53)	(-0.39)
All	-1.00	-0.64	-0.15	0.05	-0.42	-0.43	0.58
	(-1.92)	(-1.24)	(-0.24)	(0.06)	(-0.30)	(-0.76)	(0.36)
High-Low	1.44	1.71	2.62	4.26	2.68	2.55	
	(1.50)	(1.71)	(2.97)	(4.36)	(1.80)	(3.47)	
Panel B: Four-factor alpha of benchmark-adjusted return							
Active Share	Tracking error quintile						
quintile	Low	2	3	4	High	All	High-Low
High	1.44	0.79	0.48	2.72	0.29	1.15	-1.15
	(1.79)	(1.02)	(0.68)	(3.17)	(0.22)	(1.86)	(-0.74)
4	-0.11	-0.91	-0.88	-1.52	-1.64	-1.02	-1.53
	(-0.22)	(-1.17)	(-1.23)	(-1.63)	(-1.33)	(-1.63)	(-1.08)
3	-1.05	-1.41	-1.58	-2.25	-2.86	-1.83	-1.81
	(-1.97)	(-2.15)	(-2.34)	(-2.23)	(-2.51)	(-2.84)	(-1.59)
2	-1.46	-1.47	-1.82	-2.67	-3.43	-2.18	-1.97
	(-3.31)	(-2.29)	(-2.99)	(-3.31)	(-3.61)	(-4.00)	(-2.17)
Low	-1.29	-1.36	-1.66	-2.26	-2.57	-1.83	-1.28
	(-4.80)	(-4.80)	(-4.33)	(-4.43)	(-3.73)	(-5.01)	(-2.13)
All	-0.50	-0.87	-1.09	-1.20	-2.05	-1.14	-1.55
	(-1.45)	(-2.13)	(-2.58)	(-1.81)	(-2.28)	(-2.53)	(-1.68)
High-Low	2.73	2.16	2.13	4.99	2.86	2.98	
	(3.33)	(2.52)	(2.61)	(5.60)	(2.26)	(4.51)	

Table 7: **Gross equal-weighted alphas for all-equity mutual funds in 1990-2003**, sorted by the two dimensions of active management. The measures of active management are computed as before. Gross fund returns are the returns on a fund's portfolio and do not include any fees or transaction costs. Index funds are excluded from the sample. The table shows annualized returns, followed by t-statistics (in parentheses) based on White's standard errors.

Panel A: Benchmark-adjusted return

Active Share	Tracking error quintile						
quintile	Low	2	3	4	High	All	High-Low
High	1.34	1.56	3.01	3.34	2.72	2.40	1.38
	(1.61)	(1.67)	(3.30)	(2.70)	(1.29)	(2.80)	(0.60)
4	1.02	1.32	1.35	1.39	1.58	1.33	0.56
	(1.56)	(1.59)	(1.28)	(0.97)	(0.64)	(1.28)	(0.20)
3	0.09	0.78	0.97	1.19	1.06	0.81	0.97
	(0.16)	(1.03)	(1.10)	(1.00)	(0.54)	(0.94)	(0.46)
2	-0.24	0.13	0.68	0.59	0.02	0.24	0.26
	(-0.54)	(0.23)	(0.93)	(0.59)	(0.01)	(0.34)	(0.17)
Low	0.00	0.37	0.06	0.22	-0.08	0.11	-0.08
	(-0.02)	(1.24)	(0.14)	(0.40)	(-0.10)	(0.29)	(-0.10)
All	0.44	0.83	1.22	1.35	1.05	0.98	0.61
	(1.04)	(1.56)	(1.82)	(1.42)	(0.63)	(1.41)	(0.34)
High-Low	1.35	1.19	2.95	3.13	2.81	2.29	
	(1.68)	(1.27)	(3.47)	(2.79)	(1.70)	(3.05)	

Panel B: Four-factor alpha of benchmark-adjusted return

Active Share	Tracking error quintile						
quintile	Low	2	3	4	High	All	High-Low
High	1.39	0.86	1.38	2.50	1.37	1.51	-0.02
	(1.80)	(1.00)	(1.76)	(2.54)	(1.02)	(2.23)	(-0.02)
4	0.42	-0.23	-0.92	-1.28	-1.08	-0.63	-1.50
	(0.76)	(-0.27)	(-1.07)	(-1.12)	(-0.78)	(-0.81)	(-1.01)
3	-0.33	-0.38	-0.88	-1.24	-1.58	-0.89	-1.26
	(-0.55)	(-0.52)	(-1.15)	(-1.19)	(-1.40)	(-1.28)	(-1.08)
2	-0.59	-0.83	-0.60	-1.58	-2.21	-1.17	-1.63
	(-1.42)	(-1.59)	(-0.98)	(-1.93)	(-2.32)	(-2.14)	(-1.73)
Low	-0.30	-0.13	-0.67	-0.72	-1.31	-0.63	-1.01
	(-1.18)	(-0.47)	(-1.62)	(-1.41)	(-1.98)	(-1.73)	(-1.69)
All	0.12	-0.14	-0.34	-0.46	-0.97	-0.36	-1.09
	(0.32)	(-0.28)	(-0.63)	(-0.62)	(-1.06)	(-0.69)	(-1.18)
High-Low	1.69	0.99	2.05	3.22	2.68	2.13	
	(2.15)	(1.12)	(2.57)	(3.41)	(2.21)	(3.29)	

Table 8: **Net equal-weighted alphas for all-equity mutual funds in 1990-2003**, sorted by fund size and Active Share (sequentially and in that order). Active Share is computed as before. Net fund returns are the returns to a fund investor after fees and transaction costs. Index funds are excluded from the sample. The table shows annualized returns, followed by t-statistics (in parentheses) based on White's standard errors.

Panel A: Benchmark-adjusted return							
Active Share	Fund size quintile						
quintile	Low	2	3	4	High	All	High-Low
High	1.15	1.64	1.07	0.52	-0.27	0.82	-1.42
	(1.15)	(1.71)	(1.33)	(0.64)	(-0.36)	(1.20)	(-1.30)
4	1.88	1.02	1.01	-0.21	-0.61	0.62	-2.49
	(2.13)	(1.08)	(1.11)	(-0.24)	(-0.67)	(0.78)	(-3.41)
3	-0.12	-0.53	-1.16	-1.27	-0.52	-0.72	-0.40
	(-0.13)	(-0.63)	(-1.33)	(-1.49)	(-0.61)	(-0.91)	(-0.66)
2	-1.40	-2.05	-1.49	-0.76	-1.54	-1.45	-0.14
	(-1.82)	(-3.00)	(-1.84)	(-0.94)	(-2.55)	(-2.20)	(-0.26)
Low	-1.77	-1.48	-1.45	-1.20	-1.29	-1.44	0.48
	(-3.36)	(-3.16)	(-3.09)	(-2.72)	(-3.39)	(-3.48)	(1.37)
All	-0.05	-0.28	-0.40	-0.58	-0.85	-0.43	-0.79
	(-0.09)	(-0.48)	(-0.68)	(-0.92)	(-1.38)	(-0.76)	(-1.99)
High-Low	2.92	3.12	2.53	1.72	1.02	2.26	
	(2.66)	(3.03)	(2.83)	(2.29)	(1.50)	(3.20)	
Panel B: Four-factor alpha of benchmark-adjusted return							
Active Share	Fund size quintile						
quintile	Low	2	3	4	High	All	High-Low
High	1.71	1.39	1.15	0.19	-0.67	0.75	-2.39
	(1.97)	(1.58)	(1.61)	(0.25)	(-1.01)	(1.26)	(-2.60)
4	0.87	-0.24	-0.02	-1.55	-1.90	-0.57	-2.78
	(1.09)	(-0.28)	(-0.02)	(-1.72)	(-2.61)	(-0.82)	(-3.74)
3	-1.47	-1.60	-2.11	-2.58	-1.54	-1.86	-0.07
	(-2.21)	(-2.13)	(-2.97)	(-3.71)	(-2.36)	(-3.09)	(-0.12)
2	-1.95	-2.52	-2.79	-1.56	-2.16	-2.20	-0.21
	(-3.24)	(-4.27)	(-4.56)	(-2.04)	(-3.88)	(-4.02)	(-0.43)
Low	-2.06	-1.81	-1.90	-1.64	-1.69	-1.82	0.38
	(-3.97)	(-4.04)	(-4.61)	(-3.89)	(-5.30)	(-4.81)	(1.03)
All	-0.59	-0.96	-1.14	-1.43	-1.60	-1.14	-1.01
	(-1.34)	(-1.93)	(-2.41)	(-2.58)	(-3.41)	(-2.53)	(-3.05)
High-Low	3.78	3.20	3.05	1.83	1.01	2.57	
	(3.74)	(3.22)	(3.75)	(2.51)	(1.43)	(3.87)	

Table 9: **Net equal-weighted alphas for all-equity mutual funds in 1990-2003**, sorted by Active Share and prior one-year return (sequentially and in that order). The prior return on a fund is measured as its benchmark-adjusted gross return over the previous 12 months. Only funds with at least 9 months of such returns are included. Active Share is computed as before. Net fund returns are the returns to a fund investor after fees and transaction costs. Index funds are excluded from the sample. The table shows annualized returns, followed by t-statistics (in parentheses) based on White's standard errors.

Panel A: Benchmark-adjusted return							
Active Share	Prior 1-year return quintile						
quintile	Low	2	3	4	High	All	High-Low
High	-1.71	-0.51	0.74	1.32	5.10	0.98	6.81
	(-1.42)	(-0.52)	(0.94)	(1.58)	(3.67)	(1.40)	(3.35)
4	-2.16	-1.05	-0.57	1.16	3.01	0.08	5.18
	(-1.97)	(-1.17)	(-0.71)	(1.25)	(1.72)	(0.11)	(2.36)
3	-2.65	-1.81	-1.21	-0.19	1.28	-0.92	3.93
	(-2.37)	(-2.10)	(-1.42)	(-0.21)	(0.88)	(-1.17)	(2.00)
2	-2.28	-1.84	-1.80	-0.52	-0.06	-1.30	2.23
	(-2.32)	(-2.64)	(-2.62)	(-0.66)	(-0.05)	(-1.88)	(1.46)
Low	-2.08	-2.06	-1.48	-1.00	-0.39	-1.40	1.69
	(-3.67)	(-4.91)	(-3.85)	(-2.12)	(-0.53)	(-3.46)	(1.91)
All	-2.18	-1.47	-0.88	0.15	1.77	-0.52	3.95
	(-2.59)	(-2.38)	(-1.60)	(0.24)	(1.48)	(-0.92)	(2.46)
High-Low	0.37	1.55	2.23	2.32	5.50	2.39	
	(0.34)	(1.61)	(2.62)	(2.52)	(4.99)	(3.27)	
Panel B: Four-factor alpha of benchmark-adjusted return							
Active Share	Prior 1-year return quintile						
quintile	Low	2	3	4	High	All	High-Low
High	-0.98	0.39	0.89	1.04	3.50	0.96	4.48
	(-1.01)	(0.46)	(1.25)	(1.27)	(3.29)	(1.56)	(3.06)
4	-2.28	-1.83	-1.90	-0.44	0.46	-1.19	2.74
	(-2.26)	(-2.14)	(-2.84)	(-0.59)	(0.41)	(-1.90)	(1.72)
3	-2.48	-2.65	-2.08	-1.78	-0.99	-2.00	1.49
	(-2.39)	(-3.24)	(-2.67)	(-2.50)	(-1.04)	(-3.10)	(1.00)
2	-2.55	-2.16	-2.41	-1.77	-1.80	-2.14	0.75
	(-2.47)	(-3.21)	(-4.60)	(-3.14)	(-2.30)	(-3.90)	(0.59)
Low	-2.05	-2.26	-1.78	-1.52	-1.58	-1.84	0.47
	(-3.41)	(-5.85)	(-4.97)	(-3.42)	(-3.14)	(-5.01)	(0.70)
All	-2.07	-1.72	-1.47	-0.90	-0.11	-1.26	1.96
	(-2.69)	(-3.08)	(-3.30)	(-1.99)	(-0.15)	(-2.76)	(1.74)
High-Low	1.07	2.65	2.67	2.56	5.08	2.80	
	(1.15)	(3.14)	(3.45)	(2.67)	(4.91)	(4.33)	

Table 10: **Equal-weighted performance of the benchmark indexes of all-equity mutual funds in 1990-2003**, sorted by the two dimensions of active management. The measures of active management are computed as before. These are returns on official benchmark indexes and thus do not include any fees or transaction costs. Index funds are excluded from the sample. The table shows annualized returns, followed by t-statistics (in parentheses) based on White's standard errors.

Panel A: Excess return (relative to risk-free rate)							
Active Share	Tracking error quintile						
quintile	Low	2	3	4	High	All	High-Low
High	8.64	8.88	7.18	7.07	6.95	7.74	-1.69
	(2.01)	(1.93)	(1.51)	(1.36)	(1.39)	(1.64)	(-0.88)
4	8.14	8.66	7.34	6.70	7.51	7.67	-0.64
	(1.95)	(1.93)	(1.57)	(1.37)	(1.55)	(1.68)	(-0.30)
3	7.89	8.40	7.34	6.71	7.76	7.61	-0.13
	(2.02)	(2.09)	(1.67)	(1.46)	(1.69)	(1.80)	(-0.06)
2	7.54	7.89	7.38	7.28	7.38	7.50	-0.15
	(1.96)	(1.95)	(1.71)	(1.66)	(1.68)	(1.81)	(-0.09)
Low	7.19	7.62	7.60	7.72	7.13	7.46	-0.06
	(1.82)	(1.91)	(1.84)	(1.80)	(1.65)	(1.81)	(-0.06)
All	7.88	8.29	7.37	7.10	7.35	7.60	-0.53
	(2.01)	(2.01)	(1.69)	(1.56)	(1.62)	(1.78)	(-0.32)
High-Low	1.45	1.25	-0.42	-0.64	-0.18	0.29	
	(0.60)	(0.52)	(-0.17)	(-0.25)	(-0.07)	(0.12)	
Panel B: Four-factor alpha of excess return							
Active Share	Tracking error quintile						
quintile	Low	2	3	4	High	All	High-Low
High	-1.36	-0.66	-1.40	-2.00	-1.58	-1.40	-0.22
	(-1.07)	(-0.54)	(-1.16)	(-2.01)	(-1.64)	(-1.36)	(-0.20)
4	-1.28	0.57	-0.49	-0.82	0.51	-0.29	1.79
	(-1.26)	(0.67)	(-0.61)	(-0.99)	(0.67)	(-0.43)	(1.53)
3	-0.23	1.17	0.68	0.48	1.69	0.76	1.92
	(-0.31)	(1.99)	(1.16)	(0.77)	(2.33)	(1.70)	(1.61)
2	0.17	1.12	1.20	1.45	1.79	1.15	1.61
	(0.33)	(2.28)	(2.19)	(2.54)	(2.62)	(2.83)	(1.72)
Low	0.64	1.20	1.36	1.86	1.69	1.35	1.05
	(1.70)	(2.91)	(2.75)	(2.93)	(2.57)	(3.05)	(1.69)
All	-0.41	0.68	0.27	0.20	0.82	0.31	1.23
	(-0.60)	(1.19)	(0.51)	(0.36)	(1.45)	(0.66)	(1.43)
High-Low	-2.00	-1.86	-2.76	-3.85	-3.27	-2.75	
	(-1.72)	(-1.55)	(-2.22)	(-3.36)	(-2.69)	(-2.53)	

Table 11: **Predictive regression for fund performance in 1992-2003.** The dependent variable in columns 1-4 is based on the cumulative net return (after all expenses) over calendar year t, while the independent variables are measured at the end of year $t-1$. Alphas are computed with respect to the four-factor model. All explanatory variables are computed as before in Table 4. Index funds are excluded from the sample. Since the expense ratio and manager tenure are missing before 1992, we limit all specifications to the same time period. All specifications include year dummies, and columns 3-6 also include benchmark dummies. The t-statistics (in parentheses) are based on standard errors clustered by year.

	Benchmark-adjusted alphas		Excess return (over risk-free rate) alphas		Characteristic Selectivity	
	(1)	(2)	(3)	(4)	(5)	(6)
Active Share	0.0722	0.0666	0.0496	0.0450	0.0200	0.0154
	(2.53)	(2.42)	(1.89)	(1.75)	(1.66)	(1.31)
Active Share		0.0139		0.0133		0.0136
* (Below median size)		(2.36)		(2.55)		(5.07)
Tracking error	-0.1454	-0.1459	-0.1006	-0.1022	-0.0567	-0.0578
	(1.76)	(1.75)	(1.71)	(1.72)	(0.66)	(0.68)
Turnover	-0.0041	-0.0041	-0.0033	-0.0033	0.0003	0.0003
	(0.84)	(0.84)	(0.78)	(0.78)	(0.10)	(0.10)
Expenses	-1.3117	-1.3024	-1.3447	-1.3348	-0.2797	-0.2694
	(6.00)	(5.99)	(5.10)	(5.06)	(0.75)	(0.73)
log_{10}(TNA)	-0.0187	-0.0013	-0.0127	0.0040	-0.0019	0.0155
	(3.06)	(0.14)	(2.22)	(0.43)	(0.31)	(2.39)
$(log_{10}(\text{TNA}))^2$	0.0025	0.0003	0.0015	-0.0005	0.0003	-0.0019
	(2.12)	(0.22)	(1.46)	(0.40)	(0.23)	(1.60)
Number of stocks / 100	0.0043	0.0043	0.0027	0.0028	0.0014	0.0015
	(3.81)	(3.82)	(2.90)	(2.97)	(3.59)	(3.70)
Fund age / 100	-0.0213	-0.0214	-0.0129	-0.0130	-0.0088	-0.0089
	(3.45)	(3.44)	(3.67)	(3.59)	(1.80)	(1.80)
Manager tenure / 100	0.0245	0.0250	0.0199	0.0201	0.0452	0.0453
	(0.85)	(0.87)	(0.75)	(0.77)	(2.50)	(2.50)
Inflow, t-1 to t	0.0025	0.0024	-0.0018	-0.0019	0.0033	0.0033
	(0.77)	(0.74)	(0.53)	(0.56)	(0.88)	(0.85)
Inflow, t-3 to t-1	0.0010	0.0010	0.0011	0.0012	0.0012	0.0012
	(1.00)	(1.05)	(1.05)	(1.09)	(1.39)	(1.44)
Return over index, t-1 to t	0.0745	0.0739	0.0757	0.0753	0.0572	0.0566
	(1.36)	(1.35)	(1.51)	(1.50)	(1.52)	(1.51)
Return over index, t-3 to t-1	-0.0353	-0.0350	-0.0378	-0.0374	-0.0749	-0.0746
	(1.72)	(1.68)	(1.59)	(1.56)	(3.23)	(3.18)
Index return, t-1 to t	0.0369	0.0371	0.0010	0.0011	0.0325	0.0327
	(1.40)	(1.42)	(0.12)	(0.13)	(3.31)	(3.24)
Index return, t-3 to t-1	-0.0233	-0.0228	-0.0563	-0.0560	-0.0559	-0.0556
	(0.65)	(0.64)	(2.80)	(2.82)	(5.45)	(5.42)
N	8,232	8,232	8,232	8,232	6,615	6,615
R^2	0.0376	0.0395	0.1768	0.1778	0.168	0.1694

Table 12: **Predicting returns with different measures of active management in 1992-2003.** The dependent variable is the benchmark-adjusted cumulative net return (after all expenses) over calendar year t, while the independent variables are measured at the end of year $t-1$. Active Share and tracking error are computed as before. Industry-level Active Share is computed similarly to Active Share, except that it replaces individual stocks with 10 industry portfolios. The Industry Concentration Index is computed as in Kacperczyk, Sialm, and Zheng (2005), except that the benchmark index is selected by following the same methodology as elsewhere in this paper. The stock concentration index is computed just like the Industry Concentration Index, except that it uses individual stocks rather than industry portfolios. Turnover is an annualized value. The control variables include all the remaining variables in Table 11. Index funds are excluded from the sample. Since the expense ratio and manager tenure are missing before 1992, we limit all specifications to the same time period. All specifications include year dummies, and columns 9-10 also include benchmark dummies. The t-statistics (in parentheses) are based on standard errors clustered by year.

	Benchmark-adjusted alphas								Excess return (over risk-free rate) alphas	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Active Share	0.0560						0.0751	0.0612	0.0067	0.0172
	(2.70)						(2.39)	(2.11)	(0.32)	(0.73)
Active Share							0.0115	0.0141	0.0123	0.0130
* (Below median size)							(2.26)	(2.40)	(3.07)	(2.52)
Tracking error		-0.0604					-0.2017	-0.1463	-0.1870	-0.1181
		(0.62)					(1.76)	(1.77)	(2.06)	(1.88)
Industry-level Active Share			0.0396				-0.0283	0.0464	0.0303	0.0845
			(1.82)				(0.56)	(0.93)	(0.60)	(1.83)
Industry Concentration Index				0.0816			0.1321	-0.0755	0.0606	-0.1243
				(2.74)			(1.84)	(1.13)	(0.84)	(2.61)
Stock concentration Index					0.0388		-0.3070	-0.0503	0.0763	0.1434
					(0.18)		(1.18)	(0.18)	(0.31)	(0.58)
Turnover						-0.0019	-0.0021	-0.0039	-0.0007	-0.0029
						(0.35)	(0.43)	(0.78)	(0.14)	(0.69)
Control variables	No	No	No	No	No	No	No	Yes	No	Yes
N	11,480	11,480	11,480	11,480	11,481	11,351	11,351	8,232	11,351	8,232
R^2	0.0098	0.0004	0.0032	0.0031	0.0002	0.0007	0.0184	0.0397	0.1381	0.1791

Appendix C: Figures

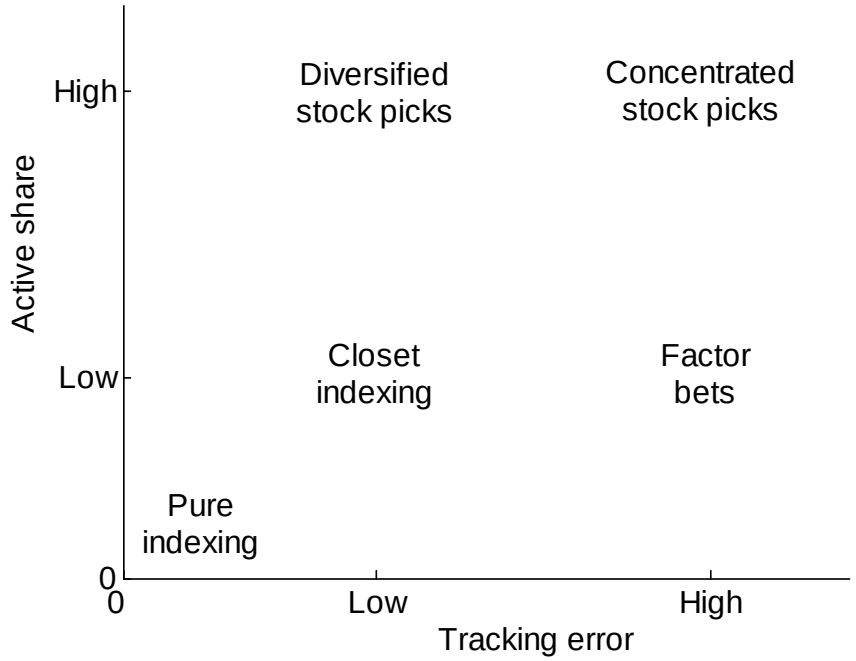

Figure 1: Different types of active and passive management, as revealed by a two-dimensional picture. Active Share represents the fraction of portfolio holdings that differ from the benchmark index, thus emphasizing stock selection. Tracking error is the volatility of fund return in excess of the benchmark, so it emphasizes bets on systematic risk.

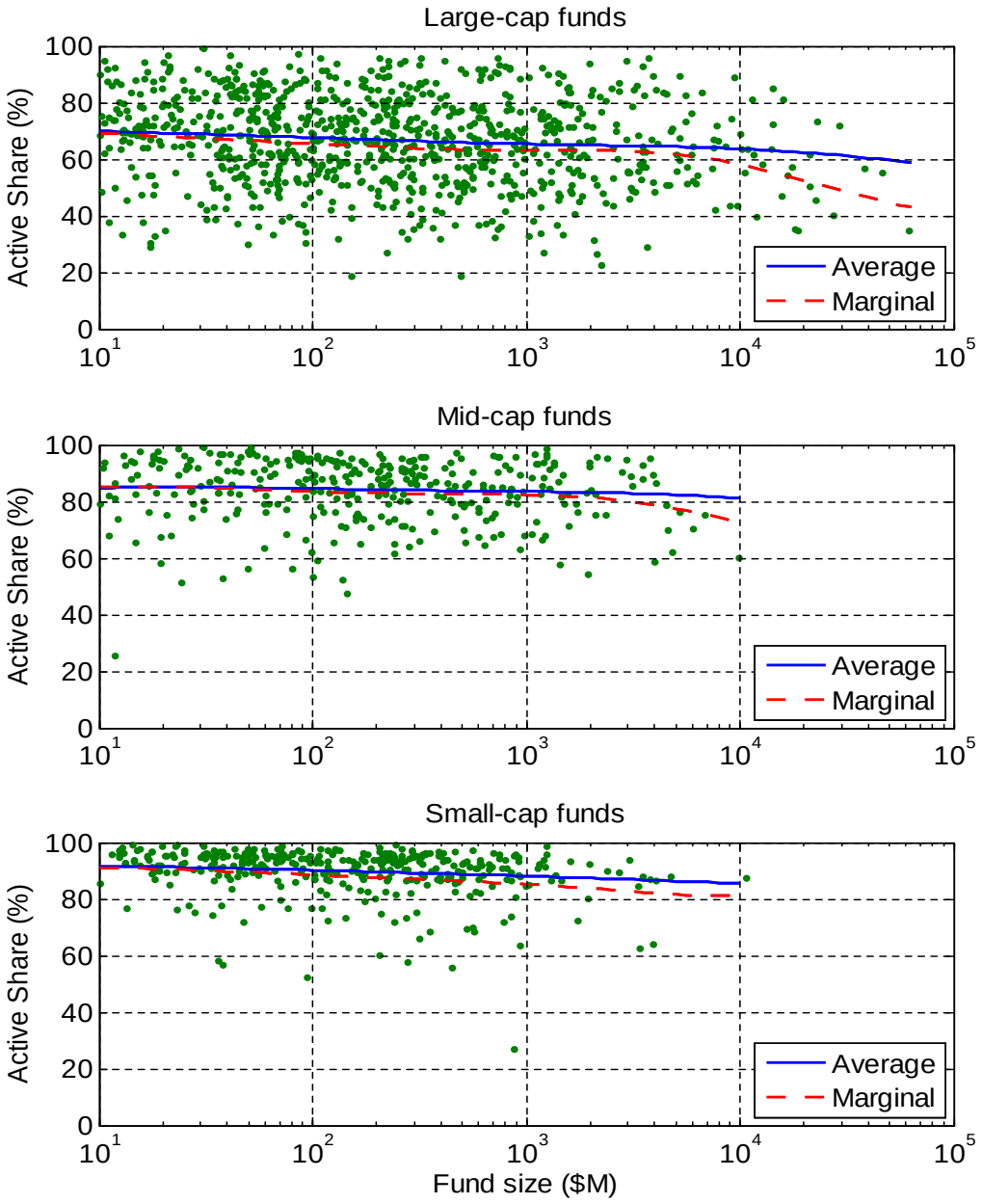

Figure 2: Average Active Share for US all-equity mutual funds and the Active Share of a marginal dollar in 2002. Fund size is total net assets expressed in millions of dollars. We exclude all index funds (Active Share less than 20%) and funds with less than $10M in stock holdings. The average Active Share is estimated from a nonparametric kernel regression with a Gaussian kernel and bandwidth equal to 0.5.

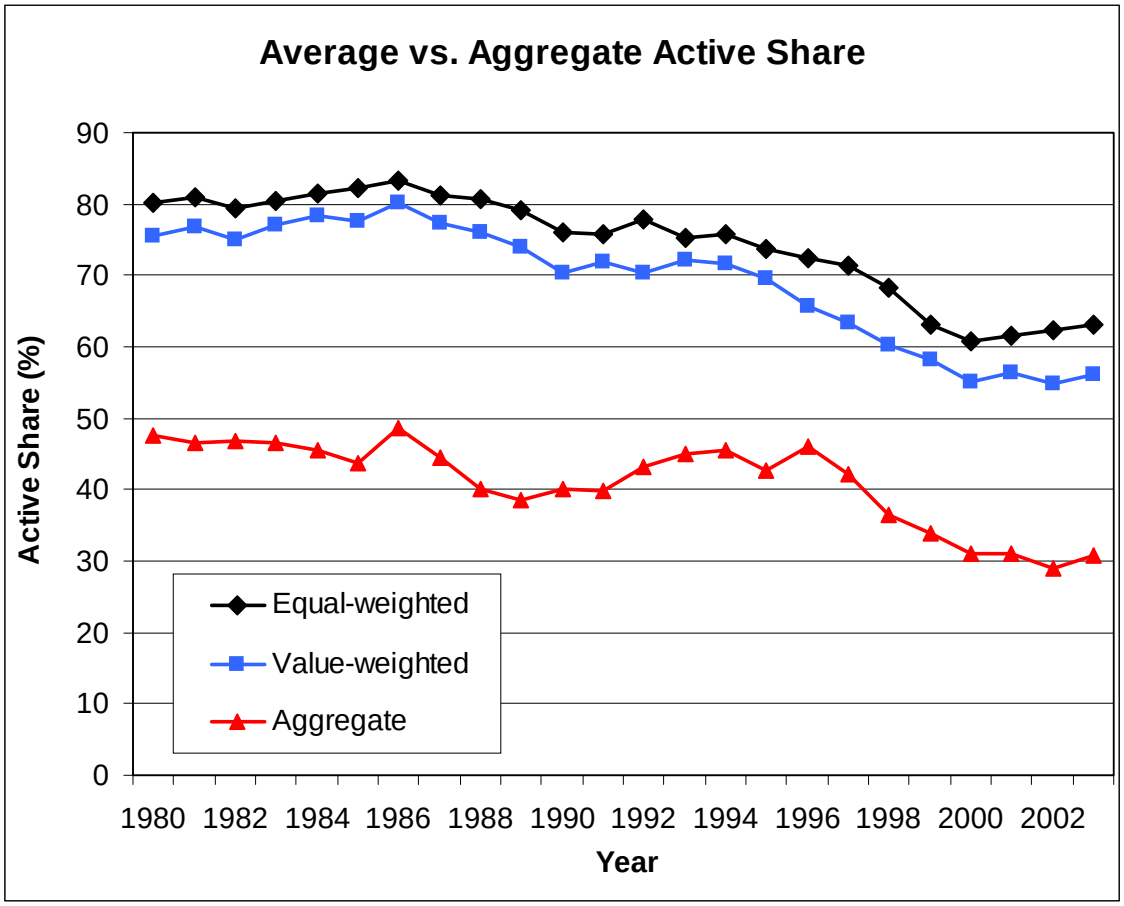

Figure 3: **Aggregate-level and fund-level Active Share** for active funds with S&P 500 as the benchmark index. Each year we compute the equal-weighted and value-weighted (by fund size) Active Share across the funds. We also aggregate the funds' portfolios into one aggregate portfolio and compute its Active Share. Index funds are excluded from the sample.