Fund Style, Characteristic-Matched Performance Benchmarks and Activity Measures: A New Approach*

Daniel Buncic
Institute of Mathematics & Statistics
University of St. Gallen
Switzerland
daniel.buncic@unisg.ch

Jon E. Eggins
Russell Investment Group
909 A Street, Tacoma
WA98402, USA
jeggins@russell.com

Robert J. Hill†
Department of Economics
University of Graz
Austria
robert.hill@uni-graz.at

October 19, 2011

Abstract

We propose a new approach for constructing characteristic-matched performance benchmarks that simultaneously generates new measures of fund style and activity. A fund’s optimal benchmark portfolio is the linear combination of two or more reference portfolios in the style space that in a least-squares sense most closely approximates the fund’s portfolio. A fund and its optimal benchmark are exactly style matched. The resulting linear combination scalar can itself be interpreted as a measure of fund style, and the distance between a fund and its optimal benchmark as a measure of fund activity. Applying our approach to a data set of US institutional fund managers, we show how some of our benchmarks generate tracking error volatilities that are comparable with the best in the literature. We also document changes in average fund style over time and shed new light on the link between activity and performance.

Keywords: Performance Measurement; Characteristic-Matched Benchmark; Size Profile; Growth Profile; Activity; Excess Return; Tracking Error

*The views expressed in this paper are those of the authors and do not necessarily reflect those of the Russell Investment Group.
†Corresponding Author
Abstract

We propose a new approach for constructing characteristic-matched performance benchmarks that simultaneously generates new measures of fund style and activity. A fund’s optimal benchmark portfolio is the linear combination of two or more reference portfolios in the style space that in a least-squares sense most closely approximates the fund’s portfolio. A fund and its optimal benchmark are exactly style matched. The resulting linear combination scalar can itself be interpreted as a measure of fund style, and the distance between a fund and its optimal benchmark as a measure of fund activity. Applying our approach to a data set of US institutional fund managers, we show how some of our benchmarks generate tracking error volatilities that are comparable with the best in the literature. We also document changes in average fund style over time and shed new light on the link between activity and performance.

Keywords: Performance Measurement; Characteristic-Matched Benchmark; Size Profile; Growth Profile; Activity; Excess Return; Tracking Error
Introduction

The measurement of style and performance of managed portfolios is of fundamental importance to the investment industry. Accurate measurement of style allows investors and funds to obtain their desired exposures to particular investment styles. It is also a prerequisite to reliable performance measurement, since the performance of a fund manager should be judged relative to an appropriate style benchmark. For example, a manager focusing on small-cap stocks should be evaluated against a small-cap benchmark, while a manager that follows a growth investing style should have performance compared to a growth benchmark.

Two main approaches have been used in the literature for ensuring that funds are matched with appropriate benchmarks. The first is regression based using past returns. Its point of departure is the capital asset pricing model (CAPM). A tailored benchmark is obtained by taking a linear combination of the risk-free and market benchmark rate of returns, with the weight for each fund determined by its value of beta. The basic model can be extended by including additional non-CAPM factors such as size, valuation and momentum (see Fama and French 1992, 1996, and Carhart 1997). Sharpe (1988, 1992) proposes an alternative regression model based on asset class factors. He uses this to determine the effective mix of a portfolio in terms of the underlying asset classes. A tailored benchmark can then be constructed as a weighted average of these asset class benchmarks, with the weights determined from the regression equation. Brown and Goetzmann (1997) propose another variant on the regression method that instead uses cluster analysis to match portfolios and benchmarks.

Regression-based methods only require data on fund performance, and factors such as size and the price-to-book ratio. To estimate the regression equation, however, a fairly long time series of observations is required. This can be problematic since the style of a fund and the factor loadings can change over time (see Annaert and Van Campenhout 2007).
The second approach requires portfolio holdings data, but has the advantage that it requires only short time horizons. Expositors of this approach include Daniel, Grinblatt, Titman and Wermers (1997), henceforth DGTW, Kothari and Warner (2001), Chan, Chen and Lakonishok (2002), henceforth CCL, Pinnuck (2003), and Chan, Dimmock and Lakonishok (2009), henceforth CDL. These authors match characteristics at the level of individual stocks. Stocks are first sorted in each style dimension, and then divided into quintile blocks. In two dimensions (say size and value-growth), this generates a total of 25 blocks. CDL show that the resulting benchmarks can be quite sensitive to the way these sorts are done (e.g., whether or not growth is sorted independently of size). DGTW construct a market-cap-weighted benchmark portfolio from each block, while CCL construct equal-weighted portfolios from each block. CDL also show that the choice between market-cap and equal weighting can significantly affect the results. Each stock in a portfolio is matched with the benchmark portfolio with the most similar style characteristics. The excess return on each stock is measured by the difference between its return and its benchmark’s return. An overall performance benchmark for a portfolio is then obtained by taking the weighted mean of these excess returns, where each stock is weighted by its dollar share of the portfolio. Each fund, therefore, has its own distinct performance benchmark.

In this paper we develop a new variant on this second approach that differs from DGTW, CCL and CDL in that our matching of characteristics is done at the level of portfolios, rather than at the level of individual stocks. This is achieved by taking linear combinations of reference portfolios in a style dimension (such as the Russell 3000 and Russell 3000 Equal-Weighted in the size dimension or the Russell 3000 Growth and Russell 3000 Value in the value-growth dimension). The optimal benchmark portfolio

---

1By assets benchmarked, Russell’s style indices account for more than 98 percent of market share for US equity growth and value oriented products (see Russell Investments 2008). Hence Russell indices are a natural source of reference style portfolios, although indices from other index providers such as Standard and Poor’s could be used instead.
for a particular fund is the linear combination of the reference portfolios that in a least-squares sense most closely approximates that fund’s portfolio. We show that this optimal benchmark achieves an exact style match in the chosen dimension with the fund’s portfolio.

The optimal value of the linear combination scalar can itself be interpreted as a measure of a fund’s style, and the distance between the portfolios of a fund and its optimal benchmark as a measure of fund activity relative to its benchmark.\(^2\)

Our approach, which has some similarities with Sharpe’s effective-mix method, simultaneously generates new characteristic-matched performance benchmarks and new measures of style and activity. It is intuitive and computationally simple.

We apply our methodology to a US institutional funds data set over the period 2002-2009. In the size and value-growth style dimensions, we illustrate the cross-section diversity of styles across managers and document the shift in style towards value stocks from 2002 to 2009.

The tracking error volatilities of at least some of our performance benchmarks are comparable with the best in the literature. This confirms the viability of our approach to benchmark construction.

We find that average style-adjusted excess returns of funds derived from both our size and value-growth characteristic-matched benchmarks are lower than those derived from the Russell 3000 benchmark. The size-adjusted excess returns, in particular, are lower since smaller stocks on average outperform larger stocks in our data set, and funds on average hold portfolios that are of smaller size than the Russell 3000 portfolio.

Finally, using our new benchmarks and activity measures, we reexamine the link between activity and performance. In contrast to the prevailing wisdom we find evidence of a negative relationship between them.

\(^2\)There is some ambiguity in the literature regarding the use of the term ‘activity’. It is sometimes used to refer to turnover. Here, however, by ‘activity’ we mean departures from passive tracking of a benchmark.
II Style Profiles and One-Dimensional Characteristic-Matched Performance Benchmarks

A Style profiles

Portfolios can be classified by style in a number of dimensions. Two of the most commonly considered styles are size and value-growth. We define a style profile $P(w)$ as a function that maps the portfolio $w$ (where $w$ is an $N \times 1$ vector defined on the $N$ assets in the portfolio) into one dimension to generate an ordinal ranking of portfolios according to that particular style. For example, suppose $P(w^1) > P(w^2)$. It follows that in this style dimension, portfolio $w^1$ attains a higher score than portfolio $w^2$.

Let $lg$ and $sm$ denote two reference portfolios in a given style dimension, such that $P(sm) < P(lg)$ (i.e., $lg$ and $sm$ are abbreviations for reference large and small portfolios in this dimension). For example, in the size dimension $lg$ could be the Russell 3000 portfolio, and $sm$ the Russell 3000 Equal-Weighted portfolio. In the value-growth dimension, $lg$ could be the Russell 3000 Growth portfolio, and $sm$ the Russell 3000 Value portfolio.

The style profile of a portfolio $w$ is measured relative to these two points of reference, which in our context are typically normalized so that $P(sm) = 0$ and $P(lg) = 1$.

B A distance minimizing benchmark portfolio and performance benchmark

Let $\tilde{w}$ denote a benchmark portfolio formed by taking linear combinations of the reference portfolios $lg$ and $sm$ as follows:

$$\tilde{w}_n = \lambda \times lg_n + (1 - \lambda) \times sm_n \quad \text{for } n = 1, \ldots, N, \quad (1)$$

where $\tilde{w}_n$ denotes the value share of stock $n$ in the benchmark portfolio, while $lg_n$ and $sm_n$ likewise denote the value shares of stock $n$ in the small and large reference
portfolios, respectively. By construction, \( \sum_{n=1}^{N} \tilde{w}_n = \sum_{n=1}^{N} lg_n = \sum_{n=1}^{N} sm_n = 1. \)

We can measure the Euclidean distance between portfolio \( w \) and \( \tilde{w} \) as follows:

\[
D = \sqrt{\sum_{n=1}^{N} (w_n - \tilde{w}_n)^2},
\]

where \( w_n \) denotes the value share of stock \( n \) in a portfolio of interest (e.g., that of a particular fund manager). Again, by construction \( \sum_{n=1}^{N} w_n = 1. \)

Substituting (1) into (2), the following expression is obtained:

\[
D = \sqrt{\sum_{n=1}^{N} [w_n - (\lambda \times lg_n + (1 - \lambda) \times sm_n)]^2}.\tag{3}
\]

Minimizing \( D \) with respect to \( \lambda \) yields the following first order condition:

\[
\frac{\partial D}{\partial \lambda} = \sum_{n=1}^{N} \left[ \frac{-lg_n + sm_n}{\sqrt{\sum_{n=1}^{N} [w_n - (\lambda \times lg_n + (1 - \lambda) \times sm_n)]^2}} \right] (w_n - \tilde{w}_n) = 0.\tag{4}
\]

Solving (4) we obtain the following solution:

\[
\hat{\lambda} = \frac{\sum_{n=1}^{N} [(lg_n - sm_n)(w_n - sm_n)]}{\sum_{n=1}^{N} (lg_n - sm_n)^2},\tag{5}
\]
and corresponding benchmark portfolio:

\[
\hat{w}_n = \hat{\lambda} \times lg_n + (1 - \hat{\lambda}) \times sm_n, \quad \text{for } n = 1, \ldots, N.\tag{6}
\]

The portfolio \( \hat{w} \) therefore is the linear combination of \( lg \) and \( sm \) that is the minimum distance from \( w \). We show in the next section that \( \hat{w} \) also achieves an exact style match with \( w \). Hence \( \hat{w} \) is an appropriate benchmark portfolio for \( w \).

Our performance benchmark for \( w \) is equal to the return on its benchmark portfolio, denoted here by \( R(\hat{w}) \). By construction, \( R(\hat{w}) = \hat{\lambda} \times R(lg) + (1 - \hat{\lambda}) R(sm) \), where \( R(lg) \) and \( R(sm) \) denote the returns on the \( lg \) and \( sm \) reference portfolios during the period of interest. If \( lg \) and \( sm \) are existing indices, such as the Russell 3000 Value and Russell 3000 Growth, then \( R(lg) \) and \( R(sm) \) can be taken directly from the index provider (in this case Russell). Meanwhile, \( \hat{\lambda} \) can be easily calculated from (5).
Our minimization problem can alternatively be expressed in regression format as follows:

\[ w_n = \lambda \times lg_n + (1 - \lambda) \times sm_n + \varepsilon_n, \]

where \( \varepsilon_n \) denotes an error term. Rearranging and switching to matrix notation yields the following:

\[ (w - sm) = \lambda(lg - sm) + \varepsilon. \]

Solving for \( \lambda \) using ordinary least squares we obtain that

\[ \hat{\lambda} = [(lg - sm)'(lg - sm)]^{-1}(lg - sm)'(w - sm), \]

which on inspection is identical to the solution for \( \hat{\lambda} \) in (5).

An analogy can be drawn here with Sharpe’s effective-mix method. We find the portfolio \( \hat{w} \) formed by taking a linear combination of the reference portfolios \( lg \) and \( sm \) that in a least-squares sense most closely approximates the portfolio \( w \) in terms of its individual asset holdings. The benchmark return is then derived as explained above from \( \hat{w} \). Sharpe by contrast calculates his benchmark return directly as the return on the linear combination of asset class factors that in a least-squares sense most closely approximates the return on portfolio \( w \). As noted above, Sharpe’s method requires a reasonably long time series to calculate its factor loadings (and assumes that the fund’s style and factor loadings are constant during this period), while our method has the advantage that given its essentially cross-section structure it can be applied over much shorter time intervals.

C  A style-profile formula derived from distance minimization

The formula for \( \hat{\lambda} \) in (5) that solves the least-squares distance minimization problem can itself be interpreted as a formula for calculating style profiles. That is, we define the following style profile:

\[ \hat{P}(w) = \frac{\sum_{n=1}^{N}[(lg_n - sm_n)(w_n - sm_n)]}{\sum_{n=1}^{N}(lg_n - sm_n)^2}. \]
The style profile $\hat{P}(w)$ has a number of attractive properties. First, from inspection of (7) it can be seen that $\hat{P}(sm) = 0$ and $\hat{P}(lg) = 1$. It follows from this that

$$\hat{P}(\hat{w}) = \hat{\lambda}\hat{P}(lg) + (1 - \hat{\lambda})\hat{P}(sm) = \hat{\lambda} = \hat{P}(w).$$

(8)

In words, (8) implies that, by construction, the distance minimizing benchmark portfolio $\hat{w}$ has the same style profile as the portfolio $w$ to which it is benchmarked. Distance minimization, therefore, here guarantees an exact style match between a portfolio and its benchmark.

The benchmark portfolio $\hat{w}$ is constructed by taking linear combinations of the two reference portfolios $lg$ and $sm$ with the relative weights on these portfolios given by the size profile $\hat{P}(w)$ as follows:

$$\hat{w}_n = \hat{P}(w)lg_n + (1 - \hat{P}(w))sm_n, \text{ for } n = 1, \ldots, N.$$  

(9)

In the size dimension, as will become apparent shortly, natural choices for $lg$ and $sm$ are market-cap and equal-weighted portfolios, respectively. Tracking funds are often defined on these portfolios. Tracking funds are likewise defined on the Russell 3000 Value and Russell Growth portfolios, which are the reference portfolios for one of our value/growth-adjusted benchmarks discussed later. These size-adjusted and value/growth-adjusted benchmarks therefore are investable (i.e., they do not incur excessive transaction costs). This is an important consideration for a performance benchmark. Huij and Verbeek (2009), for example, demonstrate that the failure to use investable benchmarks can distort the results of performance studies.

Thus far, we have derived our style-profile formula $\hat{P}(w)$ without actually providing a formal definition of style. Where desired, however, such foundations can be provided for $\hat{P}(w)$. For example, in the size dimension a natural absolute measure of style is the following:

$$S^*(w) = \sum_{n=1}^{N} w_n mc w_n,$$

(10)
where \( mcw_n \) denotes the market-cap share of stock \( n \), and \( w_n \) its value share in portfolio \( w \) as noted above. If we make the substitution \( lg_n = mcw_n \) and \( sm_n = ew_n \) (where \( ew_n = 1/N \) are the value shares of an equal-weighted portfolio) in (7), then it can be shown that our size profile \( \hat{P}^S(w) \) is a positive linear function of \( S^*(w) \) (see Appendix A), and hence achieves the same ordinal ranking of portfolios.

In the value-growth dimension, one possible absolute measure of style might be the following:

\[
G^*(w) = \prod_{n=1}^{N} \left( \frac{p_n}{b_n} \right)^{w_n},
\]

(11)

where \( p_n \) and \( b_n \) denote stock \( n \)'s price and book value per share respectively.\(^3\) Here we focus exclusively on price-to-book ratios as a measure of value. Broader measures of value that incorporate earnings, dividends or sales are considered by CDL. Our methodology can be used to construct such broader measures. We return to this issue later.

Making the substitution \( sm_n = ew_n = 1/N \) and \( lg_n = \hat{gw}_n \) in (7), where

\[
\hat{gw}_n = \frac{\ln(p_n/b_n)}{\sum_{m=1}^{N} \ln(p_m/b_m)},
\]

(12)

we obtain a growth profile \( \hat{P}^G(w) \) that is a positive monotonic function of \( G^*(w) \) and hence achieves the same ordinal ranking (again see Appendix A).

### D Some simple examples

Our method for constructing tailored performance benchmarks can be demonstrated with some examples involving a universe consisting of four stocks. Suppose 40 percent of the total market cap is in stock 1, 30 percent in stock 2, 20 percent in stock 3 and 10 percent in stock 4. One attractive feature of this formula is that its reciprocal is the market-cap-weighted geometric mean of the book-to-price ratios. That is \( 1/G^*(w) = \prod_{n=1}^{N} (b_n/p_n)^{w_n} \). Hence the ranking of portfolios does not depend on whether we focus on price-to-book or book-to-price ratios. In this latter case, the growth profile rises as one moves to the left along the growth line. This property is useful since price-to-book and book-to-price ratios contain the same information.
percent in stock 4. The holdings of nine hypothetical managers are shown in Table I. For example, Manager 1 holds only stock 1, while Manager 2 holds only stock 2.

**Insert Table I Here**

Using a market-cap portfolio as our reference $lg$ portfolio and an equal-weighted portfolio as our reference $sm$ portfolio, application of (7) yields the size profiles $\hat{P}^S(w)$ shown in Table I. Using (9), $\hat{P}^S(w^1) = 3$ implies that Manager 1’s benchmark portfolio $\hat{w}^1$ is formed by putting three times the total value invested in the market-cap portfolio and shorting two times the value invested in the equal-weighted portfolio (i.e., $\hat{w}^1_n = 3mcw_n - 2ew_n$). This yields the holdings $\hat{w}^1 = (0.7, 0.4, 0.1, -0.2)$. Our absolute size measure $S^*(w)$ defined in (10) can be used to verify that $w^1$ and $\hat{w}^1$ are the same size. $S^*(w^1) = 1 \times 0.4 = 0.4$ while $S^*(\hat{w}^1) = 0.7 \times 0.4 + 0.4 \times 0.3 + 0.1 \times 0.2 - 0.1 \times 0.2 = 0.4$.

Similarly, $\hat{P}^S(w^4) = -3$ implies that Manager 4’s benchmark portfolio $\hat{w}^4$ is formed by putting four times the total value invested in the equal-weighted portfolio and shorting three times the value invested in the market-cap-weighted portfolio (i.e., $\hat{w}^4_n = 4ew_n - 3mcw_n$), and hence $\hat{w}^4 = (-0.2, 0.1, 0.4, 0.7)$. Now $S^*(w^4) = 1 \times 0.1 = 0.1$ which is the same as $S^*(\hat{w}^4) = -0.2 \times 0.4 + 0.1 \times 0.3 + 0.4 \times 0.2 + 0.7 \times 0.2 = 0.1$.

The exact size match between $w^i$ and $\hat{w}^i$ can be demonstrated for the other Managers in Table I in an analogous manner.

### E Distance as a measure of activity

Returning to the distance measure in (2), if we replace $\hat{w}$ with $\hat{w}$, we obtain a measure $\hat{D}$ of the Euclidean distance between a portfolio $w$ and its characteristic-matched benchmark $\hat{w}$ as follows:

$$\hat{D} = \sqrt{\sum_{n=1}^{N} (w_n - \hat{w}_n)^2}. \hspace{1cm} (13)$$

We interpret $\hat{D}$ as a measure of a fund manager’s activity.footnote{4} A variant on this index

---

footnote{4}{Again it should be noted that by ‘activity’ we mean departures from passive tracking of a benchmark, rather than a measure of turnover.}
(without the square-root sign) has been used previously by Kacperczyk, Sialm and Zheng (2005), henceforth KSZ, and Brands, Brown and Gallagher (2005), henceforth BBG, in a different context. Their variant replaces $\hat{w}$ with the market portfolio. They then compare each portfolio with the market portfolio proxy (e.g., the Russell 3000), and interpret $D$ as a measure of concentration. That is, a portfolio is deemed to have zero concentration if it is identical to the market portfolio. The more it differs from the market portfolio, the more concentrated it is deemed to be relative to the market. KSZ only consider concentration over 10 industry classes, while BBG also calculate it at the level of individual stocks. Both, however, only compare portfolios with the market portfolio, and not with characteristic-matched benchmark portfolios.

In our context, $\hat{D}$ is better interpreted as a measure of the activity of a portfolio in a particular style dimension. An active portfolio can be distinguished by its deviation from its passive characteristic-matched style benchmark. Funds with a low value of $\hat{D}$ are almost style passive in that dimension in the sense that all the fund manager is effectively doing is allocating money across two passive reference portfolios. Our activity measure $\hat{D}$ is in spirit probably closest to the active share measure of Cremers and Petajisto (2009), henceforth CP. The CP activity measure differs from ours, however, in two important respects. First, it takes 19 reference portfolios (such as the S&P500, Russell 3000, Wilshire 5000), using each in turn as the benchmark and selects for a particular portfolio whichever has the lowest activity measure. In contrast, we construct benchmarks that are specifically tailored to have the same style characteristics as each portfolio, and which minimize the activity measure over a continuous multi-dimensional style space. Second, the CP activity measure optimizes using mean absolute deviation, while we use least squares.

Later in the paper, we explore the relationship between activity as we have defined it and performance.
III Two-Dimensional Characteristic-Matched Performance Benchmarks

A The need for multi-dimensional characteristic-matched style benchmarks

Fund managers may operate in two or more style dimensions, such as large cap/growth or small cap/value. In such cases, it is not enough to match portfolios and benchmarks in a single style dimension. Suppose for example that a large-cap/growth manager is evaluated against only a large-cap benchmark, and that she outperforms the benchmark. We cannot tell whether her outperformance is due to superior stock picking in the size domain (for which she should be rewarded) or due to outperformance of growth stocks (for which she should not be rewarded). A large-cap/growth manager, therefore, should be evaluated against a large-cap/growth benchmark.

Here we show how such two-dimensional characteristic-matched performance benchmarks can be constructed that simultaneously match a portfolio’s style in both the size and growth dimension.\(^5\)

To do this we now need three reference portfolios (as opposed to the two reference portfolios that we have used thus far).

B Using three reference portfolios

We distinguish between two style dimensions $S$ and $G$ (e.g., size and value-growth). The style profile of a portfolio $w$ in style dimension $S$ is now denoted by $P^S(w)$, and in dimension $G$ by $P^G(w)$. Three reference portfolios are all that are required to construct characteristic-matched benchmark portfolios in S-G space (as long as these portfolios span the space).

\(^5\)The methodology generalizes in a straightforward way to higher dimensions.
We use the notation $mcw$ to denote a market-cap-weighted portfolio such as the Russell 3000, $ew$ an equal-weighted portfolio such as the Russell 3000 Equal-Weighted, and $\hat{gw}$ as defined in (12). Substituting $mcw = lg$ and $ew = sm$ into (7), by construction, we obtain that $P^S(mcw) = 1$ and $P^S(ew) = 0$. Similarly, substituting $\hat{gw} = lg$ and $ew = sm$ in (7), we obtain that $P^G(\hat{gw}) = 1$ and $P^G(ew) = 0$. It is necessary, however, to compute the values of $P^G(mcw)$ and $P^S(\hat{gw})$.

To determine the proportions in which our three reference portfolios ($mcw$, $ew$ and $\hat{gw}$) must be combined to generate a portfolio with the same $S$ and $G$ profiles as $w$, it is necessary to solve the following system of two simultaneous equations:

$$
\begin{pmatrix}
P^S(ew) \\
\ P^G(ew)
\end{pmatrix} + \hat{\mu}_1 \begin{pmatrix}
P^S(mcw) - P^S(ew) \\
\ P^G(mcw) - P^G(ew)
\end{pmatrix} + \hat{\mu}_2 \begin{pmatrix}
P^S(\hat{gw}) - P^S(ew) \\
\ P^G(\hat{gw}) - P^G(ew)
\end{pmatrix} = \begin{pmatrix}
P^S(w) \\
\ P^G(w)
\end{pmatrix}.
$$

(14)

The only unknowns in (14) are $\hat{\mu}_1$ and $\hat{\mu}_2$. The terms $P^S(w)$ and $P^G(w)$ on the righthand side of (14) denote the size and growth profiles of the portfolio $w$ calculated from (7).

Making the substitutions $P^S(mcw) = P^G(\hat{gw}) = 1$ and $P^S(ew) = P^G(ew) = 0$, (14) simplifies to the following:

$$
\hat{\mu}_1 + \hat{\mu}_2 P^S(\hat{gw}) = P^S(w),
$$

(15)

$$
\hat{\mu}_1 P^G(mcw) + \hat{\mu}_2 = P^G(w).
$$

(16)

These equations yield the following solutions for $\hat{\mu}_1$ and $\hat{\mu}_2$:

$$
\hat{\mu}_1 = \frac{P^S(w) - P^S(\hat{gw}) P^G(w)}{1 - P^S(\hat{gw}) P^G(mcw)},
$$

(17)

$$
\hat{\mu}_2 = \frac{P^G(w) - P^G(mcw) P^S(w)}{1 - P^S(\hat{gw}) P^G(mcw)}.
$$

(18)

A characteristic-matched benchmark portfolio $\hat{w}$ that has the same $S$ and $G$ profile as the portfolio $w$ can now be derived as follows:

$$
\hat{w}_n = \hat{\mu}_1 mcw_n + \hat{\mu}_2 \hat{gw}_n + (1 - \hat{\mu}_1 - \hat{\mu}_2) ew_n, \quad \text{for } n = 1, \ldots, N,
$$

(19)
with $\hat{\mu}_1$ and $\hat{\mu}_2$ derived from (17) and (18). That $P^S(\hat{w}) = P^S(w)$ can be verified as follows:

$$P^S(\hat{w}) = \hat{\mu}_1 P^S(mcw) + \hat{\mu}_2 P^S(\hat{g}w) + (1 - \hat{\mu}_1 - \hat{\mu}_2) P^S(ew) = \hat{\mu}_1 + \hat{\mu}_2 P^S(\hat{g}w),$$

since $P^S(mcw) = 1$ and $P^S(ew) = 0$. Now substituting for $\hat{\mu}_1$ and $\hat{\mu}_2$ from (17) and (18), we obtain that:

$$P^S(\hat{w}) = \left[ \frac{P^S(w) - P^S(\hat{g}w) P^G(w)}{1 - P^S(\hat{g}w) P^G(mcw)} \right] + \left[ \frac{P^G(w) - P^G(mcw) P^S(w)}{1 - P^S(\hat{g}w) P^G(mcw)} \right] P^S(\hat{g}w) = P^S(w).$$

In the same way it can be shown that $P^G(\hat{w}) = P^G(w)$.\(^6\)

The one-dimensional least-squares optimization result for $\hat{w}$ generalizes to higher dimensions. In two dimensions, we define the distance $D$ between the portfolio $w$ and a portfolio $\hat{w}$ formed by taking linear combinations of the three reference portfolios as follows:

$$D = \sqrt{\sum_{n=1}^{N} (w_n - \hat{w}_n)^2},$$

where now

$$\hat{w}_n = \mu_1 mcw_n + \mu_2 \hat{g}w_n + (1 - \mu_1 - \mu_2)ew_n.$$ \hspace{1cm} (20)

Hence we can rewrite $D$ as follows:

$$D = \sqrt{\sum_{n=1}^{N} [w_n - \mu_1 mcw_n - \mu_2 \hat{g}w_n - (1 - \mu_1 - \mu_2)ew_n]^2}. \hspace{1cm} (21)$$

Differentiating with respect to $\mu_1$ and $\mu_2$, we obtain first order conditions which on rearrangement are identical to (15) and (16) (see Appendix B). Hence the solutions for $\mu_1$ and $\mu_2$ in this least-squares minimization problem are the same as those given in (17) and (18) above. It follows that $\hat{w}$ as defined in (19) is the linear combination of the three reference portfolios that in a least-squares sense most closely approximates the portfolio $w$, as well as having the same $S$ and $G$ profiles as $w$.

\(^6\)Although we do not do it here, this approach could also be used to construct a growth style benchmark that is matched to multiple dimensions of a portfolio’s value-growth characteristics, such as book value, earnings, dividends, or sales as recommended by CDL.
IV An Application to Fund Managers and Indices

A The data set

Our data set consists of a sample of 1183 US institutional fund managers from the Russell database covering the period 2002Q2 to 2009Q3. They are representative accounts for investment managers managing institutional portfolios, and hence include portfolios managed for pension funds, endowments, sovereign wealth funds, and consultants like Russell and Mercer. They are not mutual funds, although many of the funds represented have mutual funds. Our focus on institutional funds can be justified by the fact that they control more assets than mutual funds (see Del Guercio and Tkac 2002). The data set, while not free of selection bias since larger more successful managers are more likely to be included, is based on the published Russell Mellon performance universes that is well known in the institutional investment industry. It provides a very good representation of the opportunity set of active US equity managers available to institutional investors.

B The cross-section of fund style

A scatter plot of fund manager size and growth style profiles provides a useful indication of the range and variability of fund manager behavior. One such example is provided in Figure 1 for the 464 fund managers present in our data set in 2008Q4 (this was the quarter with the most fund managers). The reference portfolios in Figure 1 are $sm=\text{Russell 3000 Equal-Weighted}$ and $lg=\text{Russell 3000}$ in the size dimension, while $sm=\text{Russell 3000 Equal-Weighted}$ and $lg=\tilde{gw}$ as defined in (12) in the value-growth dimension.

Insert Figure 1 Here

\footnote{In any given quarter the number of managers present is rather less than this. The highest number (i.e., 464) is observed in 2008Q4.}
From Figure 1 we can see that not a single fund manager holds a portfolio with a size profile smaller than that of the Russell 3000 Equal-Weighted index, 56 out of 464 portfolios have larger size profiles than the Russell 3000 index, 63 have smaller growth profiles than the Russell 3000 Equal-Weighted index and 4 have larger growth profiles than the $g\bar{w}$ index defined in (12).

The average size profile in Figure 1 is 0.64, while the average growth profile is 0.30. The correlation coefficient between the size and growth profiles is 0.12 indicating a slight positive relationship between growth and size.

C The evolution of fund style

The evolution of fund style in the size and value-growth dimensions over time is depicted in Figure 2. In Panel A it can be seen that the average size profile (as measured by $\text{Abs}_S$ defined below in section IV.D) stays reasonably constant over the period 2002 to 2009, with the suggestion of a slight dip in size from 2007 onwards. Panel B shows a clear downward trend in the growth profile (as measured by $\text{Abs}_G$ also defined in section IV.D).

Insert Figure 2 Here

The shift of fund managers towards value apparent in Figure 2-Panel B demonstrates how the maintenance of a fixed style may require the active intervention of an investor. It may also help explain changes in the relative performance of the funds management industry as a whole.

Figures 1 and 2 taken together show in an intuitive way what fund managers are doing and how their behavior changes over time. Other approaches to addressing these issues, such as risk models and equity profiles, are computationally more intensive and not necessarily as easy to interpret.
D Description of benchmarks and empirical results

Average tracking error volatilities of characteristic-matched performance benchmarks, excess returns and activity measures for funds relative to these benchmarks are shown in Table II.

Insert Table II Here

The tracking error of a benchmark is calculated as the difference between the performance of the portfolio and the performance of its benchmark. Tracking error volatility measured by the annualized standard deviation of tracking error over a sample of quarters is often used in the literature to judge the appropriateness of benchmarks (see for example CDL and CP). A lower tracking error volatility implies that the performance differential between a portfolio and its benchmark has a larger systematic component, thus increasing the usefulness of the benchmark.

When calculating tracking error volatilities, we only consider managers for which we have at least 12 consecutive quarters of data. This reduces our sample of managers in Table II to 275.\(^8\)

The first set of results in Table II replicates the value-weight conditional sort (i.e., quarterly size, within-size, BM) method used by CDL. This is their preferred method since it outperforms benchmarks constructed from attribute-matched independent sorts of portfolios, the three-factor time-series model and cross-sectional regressions of returns on stock characteristics.\(^9\) Hence we use the CDL value-weight conditional sort method

\(^8\)The choice of the number of consecutive quarters required for inclusion is somewhat subjective. CDL for example require 16 consecutive quarters. We prefer 12 since the gains in sample size (275 instead of 164 funds) in our opinion outweighs the disadvantages of having a shorter time horizon for some managers. As a robustness check we also calculate results based on the requirement of 16, 20 and 24 consecutive quarters respectively. We find that the results for these alternatives differ only marginally from those obtained for 12 consecutive quarters.

\(^9\)CDL argue also for the use of composite value-growth measures. Our method can be easily extended in this direction by defining more than one dimension in the value-growth domain, and then matching portfolios and benchmarks by style in each dimension. We do not pursue this idea here, however,
as a point of reference with which to assess the performance of our characteristic-matched benchmarks.

Four of our characteristic-matched benchmarks are compared with the CDL benchmark in Table II. Our four benchmarks are described below. The first three are absolute benchmarks in the sense that their underlying size and growth profiles are monotonic functions of our absolute measures of size ($S^*$) and growth ($G^*$) (see Appendix A). The remaining benchmark is relative in the sense that its underlying size-growth profile is defined relative to two reference portfolios without being based on any underlying absolute measure of size or growth.

1. Abs$_S$: one-dimensional size benchmark: $lg=$Russell 3000, $sm=$Russell 3000 Equal-Weighted
2. Abs$_G$: one-dimensional growth benchmark: $lg = \hat{gw}$ (as defined in (12)), $sm=$Russell 3000 Equal-Weighted
3. Abs$_{SG}$: two-dimensional size-growth benchmark: $lg_S=$Russell 3000, $sm_S=$Russell 3000 Equal-Weighted, $lg_G = \hat{gw}$, $sm_G=$Russell 3000 Equal-Weighted
4. Rel$_G$: one-dimensional growth benchmark $lg=$Russell 3000 Growth, $sm=$Russell 3000 Value

Our absolute and relative growth benchmarks Abs$_G$ and Rel$_G$ differ in two respects. First, Abs$_G$ is a monotonic function of an absolute measure of value-growth style, while Rel$_G$ is only defined relative to two reference portfolios in the value-growth dimension. Second, Abs$_G$ tends to have a size profile only slightly larger than that of the Equal-Weighted Russell 3000 index (i.e., only slightly greater than zero), while Rel$_G$ has a size profile similar to that of the Russell 3000 index (i.e., close to 1). In other words, Abs$_G$ measures growth/value tilts relative to an equal-weighted portfolio and hence to improve comparability likewise do not consider CDL’s composite value-growth measures either.
while Rel\_G measures growth/value tilts relative to a market-cap-weighted portfolio. Some implications of this distinction are explored below.

### E Excess returns of funds

To interpret the pattern of excess returns in Table II it is first useful to consider the annualized average percentage returns on the reference Russell indices. Over the period 2002Q2-2009Q3, these are as follows: Russell 3000 = 1.53, Russell 3000 Growth = 0.49, Russell 3000 Value = 2.33, and Russell 3000 Equal Weighted = 5.27.\(^{10}\)

The average fund in our data set has an Abs\_S size profile of about 0.68 (see Figure 2 Panel A), which lies in between that of the Russell 3000 Equal-Weighted and Russell 3000 indices (which by construction have Abs\_S size profiles of 0 and 1 respectively). Since the Russell 3000 Equal-Weighted index outperforms the Russell 3000 over the period 2002Q2-2009Q3, and funds on average hold portfolios of smaller size than the Russell 3000 (i.e., Abs\_S<1), it follows that a size-adjusted performance benchmark for the average fund should be higher than the Russell 3000 benchmark. This explains why the average excess return of funds relative to the Russell 3000 index is 1.06 percent per year (not shown in Table II) but only 0.05 percent relative to Abs\_S.

The even worse performance of funds relative to Abs\_G (as evidenced by the average excess return of -0.41 per year in Table II) is at least partly explained by the \(gw\) reference portfolio’s small size profile (on average it is 0.03). This is only slightly larger than that of the Equal-Weighted Russell 3000 portfolio (which by construction has a size profile of 0). Given that fund managers on average hold portfolios with a size profile of about 0.6, there is a clear size mismatch between funds and their growth-adjusted Abs\_G benchmarks (even though their style is matched in the value-growth

---

\(^{10}\)The annual rates of return here differ slightly from those on the Russell website since we have excluded a few stocks for which we could not obtain book values which are required to calculate the Abs\_G and Abs\_SG benchmarks.
dimension). Given that small stocks significantly outperformed large stocks over our sample period it is therefore not surprising that funds on average underperformed their Abs\_G benchmarks by 0.41 percent.

Average fund performance improves dramatically from -0.41 percent to +0.78 percent when the benchmark is switched from Abs\_G to Rel\_G. This is because the Rel\_G portfolio is a linear combination of the Russell 3000 Value and Russell 3000 Growth portfolios, which are both approximately market-cap weighted. It follows that Rel\_G is also approximately market-cap weighted. Hence the size profiles of funds in our data set are on average smaller than those of their tailored Rel\_G portfolios, thus causing funds on average to outperform their Rel\_G benchmarks.

One conclusion that can be drawn from this discussion is that a benchmark that matches style only in the growth dimension is of limited use over a sample period where small stocks significantly outperform large stocks,

More meaningful results are obtained by matching style simultaneously on size and value-growth. Such matching is achieved by our Abs\_SG benchmark. According to Abs\_SG, the average excess return of funds in our data set is +0.39 percent per year. This is quite a bit lower than the CDL estimate of +1.20 percent or indeed the Russell 3000 estimate of +1.06 percent (where the latter is not a tailored benchmark). Our intuition is that a tailored benchmark should lower the average excess return below +1.06 percent (as ours do), since the average fund holds a portfolio of smaller size than the Russell 3000 portfolio, and small stocks on average outperformed large stocks.

**F Tracking error of benchmarks**

Based on median tracking error volatility, the best performer in Table I is Rel\_G, followed in order by CDL, Abs\_SG, Abs\_S, and lastly Abs\_G. The mean tracking error
volatility ranking differs only in that the order of Rel\_G and CDL is reversed.\textsuperscript{11}

The relatively high tracking error of Abs\_G is probably caused by it adjusting more over time than the fund manager portfolios themselves. This is because the price-book ratio of each stock is generally more volatile than its market-cap share. The fact that both Rel\_G and Abs\_SG have lower tracking errors than Abs\_S suggests that value-growth is a major driver of shifts in funds’ approaches to stock picking.

Finally, the results in Table II demonstrate that at least some of our methods are competitive in terms of tracking error volatility with the best of the methods considered by CDL, while at the same time being conceptually simpler, easier to compute and less data intensive. Furthermore, the fact that the CDL method uses conditional sorts on size in the value-growth dimension while our methods do not, in some sense biases the comparison against our methods.

Conditional versions of our methods could be constructed by eliminating from the reference portfolios in the value-growth dimension all stocks not held in the portfolio \( w \). The remaining stocks in the reference portfolios would then be rescaled so that their shares sum to 1. For example, the growth profile of a small-cap manager would then be calculated from reference portfolios in the value-growth dimension that are themselves by construction also small cap. The reference portfolios in the value-growth dimension therefore would themselves be tailored to each particular portfolio \( w \). This conditional approach should tend to reduce the tracking error volatilities of our methods.\textsuperscript{12}

\textsuperscript{11}Exactly the same median and mean rankings of methods are obtained when the comparison is restricted to funds present for at least 16 consecutive quarters.

\textsuperscript{12}Brands, Brown and Gallagher (2005) (BBG) draw a distinction between two aspects of active management. A manager must first decide which stocks to include in a portfolio, and, second, in what proportions to hold these stocks. BBG refer to these activities as ‘stock picking’ and ‘portfolio construction’. One feature of the conditional version of our method is that it constructs benchmarks that focus exclusively on the latter activity (i.e., portfolio construction). In this sense, our conditional method could provide a useful complement to our unconditional method and existing methods for benchmark construction.
G Activity and performance

A large literature exists on the topic of whether active fund managers on average outperform passive managers (see for example Wermers 2000). Cremers and Petajisto (2009), again henceforth CP, go further and consider whether more active managers outperform less active managers. CP distinguish between two notions of activity, which they refer to as stock selection and factor timing. Stock selection measures the deviation of a portfolio from its benchmark in a particular period. Factor timing can be measured by the tracking error volatility of managers relative to their benchmarks. CP find a positive relationship between stock selection and performance but no clear relationship between factor timing activity and performance.

Here we revisit this issue in a stock selection context using our measure of activity as defined in (13). Activity quintiles and their corresponding average excess returns calculated using our four characteristic-matched benchmarks – Abs_S, Abs_G and Abs_SG, Rel_G – and the Russell 3000 as benchmark are shown in Table III.

Insert Table III Here

In Table III, funds are sorted into quintiles by activity each quarter. The constituent funds in each quintile are updated each quarter. The average activity of each quintile across all quarters is reported in Table III, where the activity of each fund is measured relative to its own characteristic-matched benchmark using the formula in (13). The quarterly excess return for each fund relative to its own characteristic-matched benchmark is then averaged across all funds in each quintile. These quintile excess returns are then averaged across all quarters. The t-statistics for the excess returns in Table III are calculated using the Heteroscedasticity and Autocorrelation Consistent (HAC) estimator of Newey and West (1987).

\(^{13}\) We could in principle also use our approach to investigate the link between tracking error volatility and performance.

\(^{14}\) It should be noted that Table III is calculated using all the 1183 funds in the data set, while Table II uses only the 275 funds that were present for at least 12 consecutive quarters. We are able to
The results in Table III are striking in that we observe the opposite result to that obtained by CP. That is, we find that more active funds tend to perform worse than less active funds. Admittedly, in a few cases the second lowest activity quintile outperforms the lowest activity quintile or the highest activity quintile outperforms the second highest activity quintile. Also, few of the t-statistics are significant at the 5 percent significance level. Nevertheless, the general pattern in Table III is reasonably clear.

There are a number of differences between our study and that of CP that may contribute to our finding. First, there is very little overlap in our time horizons. Our data set covers the period 2002Q2 to 2009Q3, while CP’s covers the period 1990 to 2003. Second, our time horizon is shorter and includes the financial crisis that started in 2007. Third, our data set consists of institutional fund managers as opposed to mutual-fund managers. Hence the lack of overlap in our samples applies to the fund managers as well as the time horizon. Fourth, our performance benchmarks are matched in terms of style to each fund manager, while CP achieve only an approximate match by searching over 19 well-known indices to find the one that minimizes their measure of activity and assigning this as the benchmark for that particular manager in that particular period.

To determine whether the financial crisis is influencing our results, we try restricting the time span of our data set to 2002Q2-2007Q1. As shown in Table III, excluding the financial crisis makes the inverse relationship between performance and activity if anything even stronger than before. Replacing our tailored benchmarks with the Russell 3000 benchmark also does not change the general thrust of our results.

Our finding therefore suggests that the relationship between activity and performance may be dependent on the time horizon or sample of fund managers and hence is perhaps more complex than previously thought.

include all funds in Table III since we are not computing tracking error volatilities. One implication of this is that the excess return and activity results in Tables II and III are not directly comparable. Also, the quarterly excess returns in Table III have not been annualized.
V Conclusion

Characteristic-matched performance benchmarks obtained from portfolio holdings data are typically constructed using a bottom-up approach which first matches individual stocks to one of a number of discrete portfolios with similar style characteristics. The overall benchmark is then calculated by taking a weighted average of the excess returns on each of the individual stocks. We have proposed here an alternative methodology that avoids this bottom-up approach and generates exact style matches between portfolios and benchmarks that minimize the least-squares distance between a fund’s portfolio and a linear combination of reference portfolios in that style dimension.

Our approach opens a new direction for research on benchmarking of fund performance. It generates new characteristic-matched performance benchmarks that are simple and intuitive, and have tracking error volatilities that are comparable with the best in the literature. In the process we also generate new measures of fund style and activity that provide new insights into the evolution of fund style over time and the impact of activity on performance.
References


Appendix

A Constructing Size/Growth Profiles that are Monotonic Functions of Absolute Measures of Size/Growth

Let $mcw$ and $ew$ denote respectively a market-cap-weighted and equal-weighted portfolio. Setting $lg = mcw$ and $sm = ew$ in (7), we obtain that

$$P(w) = \frac{\sum_{n=1}^{N}[(mcw_n - ew_n)(w_n - ew_n)]}{\sum_{n=1}^{N}(mcw_n - ew_n)^2},$$

(A.22)

where

$$mcw_n = \frac{p_nq_n}{\sum_{m=1}^{N}p_mq_m}, \quad ew_n = \frac{1}{N}, \quad \text{for } n = 1, \ldots, N.$$  

By construction, $\sum_{n=1}^{N} mcw_n = \sum_{n=1}^{N} ew_n = 1$. In what follows it is assumed that there exist at least two stocks for which $ew_n \neq mcw_n$. Otherwise the style profile $P(w)$ below is not defined.

In this case $P(w)$ is a monotonic (linear) function of the absolute size measure $S^*(w)$ as defined in (10). This can be demonstrated as follows:

$$P(w) = \frac{\sum_{n=1}^{N}[(mcw_n - ew_n)(w_n - ew_n)]}{\sum_{n=1}^{N}(mcw_n - ew_n)^2} = \frac{\sum_{n=1}^{N}(w_nmcw_n) - 1/N}{\sum_{n=1}^{N}(mcw_n)^2 - 1/N}$$

$$= \frac{S^*(w) - 1/N}{\sum_{n=1}^{N}(mcw_n)^2 - 1/N}.$$  

(A.23)

As long as the same list of stocks is used when computing $P(w)$ for all portfolios, then $\sum_{n=1}^{N}(mcw_n)^2$ and $N$ are constants since they do not depend on $w_n$. The term $\sum_{n=1}^{N}(mcw_n)^2$ is the Herfindahl-Hirschman index. It must take a value greater than or equal to $1/N$. The term $\sum_{n=1}^{N}(mcw_n)^2 - 1/N$ can be interpreted as a normalized version of the Herfindahl-Hirschman index, where its minimum value is rescaled to zero rather than $1/N$. In the special case where $\sum_{n=1}^{N}(mcw_n)^2 = 1/N$, there is no size line (since all portfolios have the same size) and $P(w)$ is not defined. This special case aside, $P(w)$ is an increasing linear (and hence monotonic) function of $S^*(w)$.
Setting \( lg = \hat{g}w \) and \( sm = ew \) in (7), where \( \hat{g}w \) is the growth weighted portfolio defined in (12), and now assuming there exist at least two stocks for which \( ew_n \neq \hat{g}w_n \), we obtain that

\[
P(w) = \frac{\sum_{n=1}^{N} [(\hat{g}w_n - ew_n)(w_n - ew_n)]}{\sum_{n=1}^{N} (\hat{g}w_n - ew_n)^2} = \frac{\sum_{n=1}^{N} (w_n\hat{g}w_n) - 1/N}{\sum_{n=1}^{N} (\hat{g}w_n^2) - 1/N} = \frac{\ln[G^*(w)]/[\sum_{n=1}^{N} \ln(p_n/b_n)] - 1/N}{\sum_{n=1}^{N} (\hat{g}w_n^2) - 1/N}.
\]

(A.24)

\( G^*(w) \) is the absolute growth measure defined in (11). As long as the same list of stocks is used when computing \( P(w) \) for all portfolios, the terms \( \sum_{n=1}^{N} \ln(p_n/b_n) \), \( N \) and \( \sum_{n=1}^{N} (\hat{g}w_n)^2 \) are constants (i.e., they do not depend on \( w_n \)). Hence \( P(w) \) is, in this case, a positive monotonic function of \( G^*(w) \).

The growth variant of the Herfindahl-Hirschman index \( \sum_{n=1}^{N} (\hat{g}w_n)^2 \) must be greater than \( 1/N \), except in the special case where all stocks have the same price-to-book ratio. In this case, all portfolios have the same growth profile and hence there is no growth line. Also, the growth weights \( gw_n \) will be negative for stocks with a price-to-book value of less than one. The only complication arising out of negative weights is that it is theoretically possible that \( [\sum_{n=1}^{N} (\hat{g}w_n)^2] - 1/N \) may be negative. In this case \( P(w) \) is a monotonically decreasing rather than increasing function of \( G^*(w) \).

B Determination that \( \hat{w} \) is the Linear Combination of Three Reference Portfolios in Two Style Dimensions that in a Least-Squares Sense Most Closely Approximates the Portfolio \( w \).

Differentiating (21) with respect to \( \mu_1 \) and \( \mu_2 \) generates the following first order conditions:

\[
\frac{\partial D}{\partial \mu_1} = \frac{\sum_{n=1}^{N} (ew_n - mcw_n)[w_n - (1 - \mu_1 - \mu_2)ew_n - \mu_1 mcw_n - \mu_2 gw_n]}{\sum_{n=1}^{N} [w_n - \mu_1 mcw_n - \mu_2 gw_n - (1 - \mu_1 - \mu_2)ew_n]^2} = 0,
\]

27
\[
\frac{\partial D}{\partial \mu_2} = \frac{\sum_{n=1}^{N} (e w_n - g w_n) (w_n - (1 - \mu_1 - \mu_2) e w_n - \mu_1 m c w_n - \mu_2 g w_n)}{\sum_{n=1}^{N} (w_n - \mu_1 m c w_n - \mu_2 g w_n - (1 - \mu_1 - \mu_2) e w_n)^2} = 0.
\]

These first order conditions can be rearranged as follows:

\[
\begin{align*}
\sum_{n=1}^{N} (e w_n - m c w_n) (w_n - e w_n) + \mu_1 \sum_{n=1}^{N} (e w_n - m c w_n) (e w_n - m c w_n) \\
\quad + \mu_2 \sum_{n=1}^{N} (e w_n - m c w_n) (e w_n - g w_n) &= 0, \\
\sum_{n=1}^{N} (e w_n - g w_n) (w_n - e w_n) + \mu_1 \sum_{n=1}^{N} (e w_n - m c w_n) (e w_n - m c w_n) \\
\quad + \mu_2 \sum_{n=1}^{N} (e w_n - m c w_n) (e w_n - g w_n) &= 0.
\end{align*}
\]

Dividing through the first equation by \(\sum_{n=1}^{N} (m c w_n - e w_n)^2\) and the second equation by \(\sum_{n=1}^{N} (g w_n - e w_n)^2\) we obtain that

\[
\begin{align*}
P^S(w) - \mu_1 - \mu_2 \frac{\sum_{n=1}^{N} (m c w_n - e w_n) (g w_n - e w_n)}{\sum_{n=1}^{N} (g w_n - e w_n)^2} &= 0, \\
P^G(w) - \mu_1 \frac{\sum_{n=1}^{N} (m c w_n - e w_n) (g w_n - e w_n)}{\sum_{n=1}^{N} (g w_n - e w_n)^2} - \mu_2 &= 0.
\end{align*}
\]

Noticing now that \(\sum_{n=1}^{N} (m c w_n - e w_n) (g w_n - e w_n) / \sum_{n=1}^{N} (g w_n - e w_n)^2 = P^S(gw)\) and \(\sum_{n=1}^{N} (m c w_n - e w_n) (g w_n - e w_n) / \sum_{n=1}^{N} (g w_n - e w_n)^2 = P^G(mcw)\), the first order conditions reduce to the following:

\[
\begin{align*}
P^S(w) - \mu_1 - \mu_2 P^S(gw) &= 0, \quad (B.25) \\
P^G(w) - \mu_1 P^G(mcw) - \mu_2 &= 0. \quad (B.26)
\end{align*}
\]

On inspection it can be seen that (B.25) is identical to (15) while (B.26) is identical to (16).
### Table I. Simple Examples Consisting for Four Stocks and Nine Fund Managers

<table>
<thead>
<tr>
<th></th>
<th>$mcw$</th>
<th>$lg$</th>
<th>$sm$</th>
<th>$w^1$</th>
<th>$w^2$</th>
<th>$w^3$</th>
<th>$w^4$</th>
<th>$w^5$</th>
<th>$w^6$</th>
<th>$w^7$</th>
<th>$w^8$</th>
<th>$w^9$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stock 1</td>
<td>0.4</td>
<td>0.4</td>
<td>0.25</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
<td>0.33</td>
<td>0.33</td>
<td>0</td>
<td>0.4</td>
</tr>
<tr>
<td>Stock 2</td>
<td>0.3</td>
<td>0.3</td>
<td>0.25</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.33</td>
<td>0.33</td>
<td>0.6</td>
<td>0</td>
</tr>
<tr>
<td>Stock 3</td>
<td>0.2</td>
<td>0.2</td>
<td>0.25</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.33</td>
<td>0.4</td>
<td>0.6</td>
</tr>
<tr>
<td>Stock 4</td>
<td>0.1</td>
<td>0.1</td>
<td>0.25</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.5</td>
<td>0.33</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

$\hat{P}_S(w)$: 

|          | 1     | 0     | 3     | 1     | -1    | -3    | 0     | 0.33  | 1     | 0.2   | 0.6   |
|----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|

$S^*(w)$: 

|          | 0.3   | 0.25  | 0.4   | 0.3   | 0.2   | 0.1   | 0.25  | 0.27  | 0.3   | 0.26  | 0.28  |

In these examples there are four stocks. $mcw$ denotes the market-cap weights of the stocks. Stock 1 has 40 percent of the market cap, stock 2 has 30 percent, stock 3 20 percent and stock 4 the remaining 10 percent. The $lg$ reference portfolio is market-cap weighted, while the $sm$ reference portfolio is equal weighted. The portfolio holdings of Manager $i$ are shown in the $w^i$ column. Manager 1 holds only stock 1. Manager 2 holds only stock 2, etc. The size profiles $\hat{P}_S(w)$ of the manager portfolios are calculated using the formula in (7). $S^*(w)$ is the absolute measure of size of a portfolio as defined in (10). It is shown in Appendix A that $\hat{P}_S(w)$ is a positive linear function of $S^*(w)$. The benchmark portfolios $\hat{w}^i$ for each portfolio $w^i$ are obtained by inserting the size profiles $\hat{P}_S(w^i)$ into (9).
Table II: Benchmark Tracking Error Volatilities, Excess Returns and Activity for Funds (2002Q2-2009Q3)

<table>
<thead>
<tr>
<th>TE Volatility</th>
<th>CDL</th>
<th>Abs_S</th>
<th>Abs_G</th>
<th>Abs_SG</th>
<th>Rel_G</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>4.8588</td>
<td>5.7604</td>
<td>7.7792</td>
<td>5.0986</td>
<td>4.9016</td>
</tr>
<tr>
<td>Median</td>
<td>4.4690</td>
<td>5.5101</td>
<td>7.5945</td>
<td>4.8505</td>
<td>4.4663</td>
</tr>
<tr>
<td>Stdev.</td>
<td>1.7807</td>
<td>2.1238</td>
<td>1.5361</td>
<td>1.7152</td>
<td>1.9764</td>
</tr>
<tr>
<td>Min</td>
<td>1.2717</td>
<td>2.1075</td>
<td>4.4849</td>
<td>2.1481</td>
<td>1.3451</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Excess Return (%)</th>
<th>CDL</th>
<th>Abs_S</th>
<th>Abs_G</th>
<th>Abs_SG</th>
<th>Rel_G</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>1.1951</td>
<td>0.0543</td>
<td>-0.4050</td>
<td>0.3926</td>
<td>0.7795</td>
</tr>
<tr>
<td>Median</td>
<td>1.4247</td>
<td>0.3795</td>
<td>-0.2436</td>
<td>0.7309</td>
<td>0.8499</td>
</tr>
<tr>
<td>Stdev.</td>
<td>2.4661</td>
<td>2.5644</td>
<td>2.6829</td>
<td>2.5288</td>
<td>2.5497</td>
</tr>
<tr>
<td>Max</td>
<td>9.4969</td>
<td>6.4551</td>
<td>6.9226</td>
<td>7.7626</td>
<td>7.4485</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Activity</th>
<th>CDL</th>
<th>Abs_S</th>
<th>Abs_G</th>
<th>Abs_SG</th>
<th>Rel_G</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>-</td>
<td>0.1353</td>
<td>0.1454</td>
<td>0.1352</td>
<td>0.1360</td>
</tr>
<tr>
<td>Median</td>
<td>-</td>
<td>0.1302</td>
<td>0.1423</td>
<td>0.1302</td>
<td>0.1306</td>
</tr>
<tr>
<td>Stdev.</td>
<td>-</td>
<td>0.0334</td>
<td>0.0335</td>
<td>0.0333</td>
<td>0.0313</td>
</tr>
<tr>
<td>Min</td>
<td>-</td>
<td>0.0791</td>
<td>0.0833</td>
<td>0.0789</td>
<td>0.0912</td>
</tr>
<tr>
<td>Max</td>
<td>-</td>
<td>0.3223</td>
<td>0.3229</td>
<td>0.3222</td>
<td>0.3275</td>
</tr>
</tbody>
</table>

To allow the calculation of tracking error volatility, only funds that were present for at least 12 consecutive quarters are included (of which there are 264). At the beginning of each quarter each fund is matched with a characteristic-matched performance benchmark portfolio. A fund's tracking error volatility is the annualized standard deviation of the time series of quarterly differences between the fund's return and its characteristic-matched benchmark's return. A fund's excess return is the difference between the annualized percentage return on a fund and its benchmark. A fund's activity is measured by the Euclidean distance between its portfolio holdings and the portfolio holdings of its benchmark. For each method of constructing characteristic-matched performance benchmarks, the arithmetic mean, median, standard deviation, maximum and minimum of the tracking error volatilities, excess returns and activity across the 264 funds in the sample are provided. The CDL method replicates a method used by CDL which constructs characteristic-matched benchmark portfolios from 25 control portfolios from sorts first by size, and then within each size category, by book-to-market ratio. The Abs_S performance benchmark portfolio is calculated by taking a linear combination of the Equal-Weighted Russell 3000 and Russell 3000 portfolios. The Abs_G portfolio is a linear combination of the Equal-Weighted Russell 3000 and gw portfolios. The Abs_SG portfolio is a linear combination of the Equal-Weighted Russell 3000, Russell 3000 and gw portfolios. The Rel_G portfolio is a linear combination of the Russell 3000 Value and Russell 3000 Growth portfolios.
### Table III: Activity versus Performance

#### 2002Q2-2009Q3 Activity Quintiles

<table>
<thead>
<tr>
<th>Fund</th>
<th>Activity</th>
<th>Excess Return (%)</th>
<th>t-stat</th>
<th>Activity</th>
<th>Excess Return (%)</th>
<th>t-stat</th>
<th>Activity</th>
<th>Excess Return (%)</th>
<th>t-stat</th>
<th>Activity</th>
<th>Excess Return (%)</th>
<th>t-stat</th>
<th>Activity</th>
<th>Excess Return (%)</th>
<th>t-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abs_S</td>
<td>0.0988</td>
<td>0.1202</td>
<td>0.1360</td>
<td>0.1536</td>
<td>0.1951</td>
<td>0.2929</td>
<td>0.0221</td>
<td>-0.0385</td>
<td>-0.1373</td>
<td>-0.2504</td>
<td>-0.3150</td>
<td>0.1202</td>
<td>0.1360</td>
<td>0.1536</td>
<td>0.1951</td>
</tr>
<tr>
<td>Abs_G</td>
<td>0.1076</td>
<td>0.1313</td>
<td>0.1476</td>
<td>0.1650</td>
<td>0.2042</td>
<td>0.1925</td>
<td>-0.2839</td>
<td>-0.3633</td>
<td>-0.4916</td>
<td>-0.5085</td>
<td>-0.4764</td>
<td>0.1313</td>
<td>0.1476</td>
<td>0.1650</td>
<td>0.2042</td>
</tr>
<tr>
<td>Abs_SG</td>
<td>0.0988</td>
<td>0.1200</td>
<td>0.1358</td>
<td>0.1534</td>
<td>0.1948</td>
<td>0.0636</td>
<td>0.0530</td>
<td>0.0643</td>
<td>-0.0066</td>
<td>-0.0321</td>
<td>-0.0106</td>
<td>0.0636</td>
<td>0.0530</td>
<td>0.0643</td>
<td>-0.0066</td>
</tr>
<tr>
<td>Rel_SG</td>
<td>0.1032</td>
<td>0.1205</td>
<td>0.1355</td>
<td>0.1531</td>
<td>0.1971</td>
<td>0.1935</td>
<td>0.2792</td>
<td>0.2862</td>
<td>0.1938</td>
<td>0.0939</td>
<td>0.0857</td>
<td>0.1935</td>
<td>0.2792</td>
<td>0.2862</td>
<td>0.1938</td>
</tr>
<tr>
<td>R3000</td>
<td>0.1051</td>
<td>0.1242</td>
<td>0.1390</td>
<td>0.1565</td>
<td>0.1978</td>
<td>0.2528</td>
<td>0.4146</td>
<td>0.4343</td>
<td>0.2692</td>
<td>0.1264</td>
<td>0.1618</td>
<td>0.2528</td>
<td>0.4146</td>
<td>0.4343</td>
<td>0.2692</td>
</tr>
</tbody>
</table>

#### 2002Q2-2007Q1 Activity Quintiles

<table>
<thead>
<tr>
<th>Fund</th>
<th>Activity</th>
<th>Excess Return (%)</th>
<th>t-stat</th>
<th>Activity</th>
<th>Excess Return (%)</th>
<th>t-stat</th>
<th>Activity</th>
<th>Excess Return (%)</th>
<th>t-stat</th>
<th>Activity</th>
<th>Excess Return (%)</th>
<th>t-stat</th>
<th>Activity</th>
<th>Excess Return (%)</th>
<th>t-stat</th>
<th>Activity</th>
<th>Excess Return (%)</th>
<th>t-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abs_S</td>
<td>0.0994</td>
<td>0.1205</td>
<td>0.1362</td>
<td>0.1532</td>
<td>0.1957</td>
<td>0.5190</td>
<td>0.1039</td>
<td>0.0267</td>
<td>-0.1022</td>
<td>-0.2171</td>
<td>-0.4152</td>
<td>0.1039</td>
<td>0.0267</td>
<td>-0.1022</td>
<td>-0.2171</td>
<td>-0.4152</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abs_G</td>
<td>0.1091</td>
<td>0.1321</td>
<td>0.1486</td>
<td>0.1655</td>
<td>0.2055</td>
<td>0.5177</td>
<td>-0.2256</td>
<td>-0.4248</td>
<td>-0.5803</td>
<td>-0.6476</td>
<td>-0.7433</td>
<td>0.1091</td>
<td>0.1321</td>
<td>0.1486</td>
<td>0.1655</td>
<td>0.2055</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abs_SG</td>
<td>0.0994</td>
<td>0.1203</td>
<td>0.136</td>
<td>0.153</td>
<td>0.1954</td>
<td>0.3419</td>
<td>0.1863</td>
<td>0.0618</td>
<td>0.0602</td>
<td>-0.031</td>
<td>-0.1555</td>
<td>0.3419</td>
<td>0.1863</td>
<td>0.0618</td>
<td>0.0602</td>
<td>-0.031</td>
<td>-0.1555</td>
<td></td>
</tr>
<tr>
<td>Rel_SG</td>
<td>0.1034</td>
<td>0.1207</td>
<td>0.1353</td>
<td>0.1529</td>
<td>0.1987</td>
<td>0.1811</td>
<td>0.6746</td>
<td>0.2995</td>
<td>0.3779</td>
<td>-0.2249</td>
<td>-1.1044</td>
<td>0.1811</td>
<td>0.6746</td>
<td>0.2995</td>
<td>0.3779</td>
<td>-0.2249</td>
<td>-1.1044</td>
<td></td>
</tr>
<tr>
<td>R3000</td>
<td>0.1050</td>
<td>0.1244</td>
<td>0.1390</td>
<td>0.1562</td>
<td>0.1985</td>
<td>0.4115</td>
<td>0.4660</td>
<td>0.5047</td>
<td>0.2710</td>
<td>0.0751</td>
<td>0.0544</td>
<td>0.4115</td>
<td>0.4660</td>
<td>0.5047</td>
<td>0.2710</td>
<td>0.0751</td>
<td>0.0544</td>
<td></td>
</tr>
</tbody>
</table>

Funds are sorted by activity each quarter. A fund's activity is measured as the Euclidean distance between its portfolio holdings and the portfolio holdings of its tailored benchmark. The constituent funds in each quintile are updated each quarter. The activity of each quintile is then averaged across all quarters. The quarterly percentage excess return for each fund relative to its own tailored benchmark is then averaged across all funds in each quintile. These quintile excess returns are then averaged across all quarters. The t-statistics for the excess returns are calculated using the Heteroskedasticity and Autocorrelation Consistent (HAC) estimator of Newey and West (1987). Four different ways of constructing tailored benchmarks are considered. These are described in the Notes to Table II. In addition, activity measures and excess returns by quintile are also calculated using the Russell 3000 as the benchmark. The benchmark in this latter case is not tailored to each individual fund.
This figure depicts a cross-section plot of size profiles and growth profiles of all 464 fund managers in our data set in 2008Q4. Most managers have a positive growth profile (implying a tilt towards growth stocks), and size profiles between 0 and 1 (implying a size larger than equal weighting but smaller than market cap weighting). The correlation between the size and growth profiles is 0.12, suggesting a weak positive relationship between size and growth.
Panel A shows how the average size profile (Abs_S) of funds in our data set was reasonably stable over the 2002Q2-2009Q3 period. Panel B shows how the average growth profile (Abs_G) has a clear downward trend, implying a strong shift from growth to value stocks.