ACCOUNTING & FINANCE

afaanz

Accounting and Finance

Measuring fund style, performance and activity: a new style-profiling approach

Daniel Buncic[a], Jon E. Eggins[b], Robert J. Hill[c]

[a]*Institute of Mathematics & Statistics, University of St. Gallen, St. Gallen, Switzerland*
[b]*Russell Investment Group, Tacoma, WA, USA*
[c]*Department of Economics, University of Graz, Graz, Austria*

Abstract

We construct new measures of fund style, performance and activity from linear combinations of off-the-shelf stock-market indices. A fund's benchmark portfolio is a linear combination of two or more reference portfolios that in a least-squares sense most closely approximates the fund's portfolio. The resulting linear combination scalar is itself a measure of fund style and the distance between a fund and its benchmark is a measure of fund activity. Our approach has a number of advantages over existing characteristic-matching methods. We illustrate our approach using a data set of US institutional funds.

Key words: Characteristic-matched benchmark; Fund activity; Investment performance; Investment style; Portfolio management

JEL classification: G11, G23

doi: 10.1111/acfi.12047

1. 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, as the performance of a fund manager should be judged relative to an

Hill acknowledges funding from the Australian Research Council (ARC Discovery grant DP0666798). The views expressed in this paper are those of the authors and do not necessarily reflect those of the Russell Investment Group. We thank two anonymous referees for their insightful comments. Address correspondence to Robert Hill.

Received 5 September 2012; accepted 13 August 2013 by David Gallagher (Editor).

appropriate style benchmark. For example, a manager focusing on small-cap stocks should be evaluated against a small-cap benchmark, while a manager who 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 market return scaled by the fund's beta. The basic model can be extended by including additional non-CAPM factors such as size, valuation and momentum (Fama and French *et al.*, 1992, 1996; 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 factor returns that are readily available, such as size, momentum 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 as the style of a fund and the factor loadings can change over time (Annaert and Van Campenhout, 2007).

The second approach needs portfolio holdings data, but has the advantage that it can be applied over short time horizons and can better track time varying changes in factor exposures. Expositors of this approach include Daniel *et al.* (1997), henceforth DGTW, Kothari and Warner (2001), Chan *et al.* (2002), Fong *et al.* (2008) and Chan *et al.* (2009b), henceforth CDL. These authors match characteristics at the level of individual stocks. Stocks are first sorted in each style dimension. CDL, for example, then divide the universe of stocks into quintiles bins. In two dimensions (say size and value-growth), this generates a total of 25 bins. 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 bin, while Chan *et al.* (2002) construct equal-weighted portfolios from each bin. 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 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. These can often be taken off-the-shelf. Examples include the Russell 3000 and Russell 3000 Equal-Weighted in the size dimension and the Russell 3000 Growth and Russell 3000 Value in the value-growth dimension).[1] The benchmark portfolio 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. 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.

Our approach, which is similar in spirit to Sharpe's effective-mix method but using holdings rather than returns, has a number of advantages over existing characteristic-matching methods. First, it simultaneously also generates measures of a fund's style and activity. These style and activity measures are of interest in their own right.

Second, our benchmarks can often be derived by taking linear combinations of existing indices.[2] As a result, they may have relatively low associated transaction costs, particularly when the reference indices have tracking funds. This is an important consideration for a performance benchmark (Bailey, 1992). Huij and Verbeek (2009), for example, demonstrate that the failure to use investable benchmarks can distort the results of performance studies. It is also more convenient to use existing indices than to build passive characteristic-matched benchmarks that require trading.

Third, our performance benchmarks exactly style match funds and benchmarks, while for standard characteristics-matching methods the matching – which is done at the level of individual stocks – is only approximate. For example, CDL match each stock with one of 25 subsets of the stock universe in their two-dimensional style space.

Fourth, our performance benchmarks avoid the 'curse of dimensionality' problem. For example, the standard characteristic-matching method divides

---

[1] By assets benchmarked, Russell's style indices account for more than 98 per cent of market share for US equity growth and value oriented products (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.

[2] An analogy can be drawn here with Faff (2003) who uses off-the-shelf indices to construct proxies for the SMB and HML factors in the Fama-French 3-factor model.

stocks into five bins in each style dimension. In two dimensions, this requires a total of 25 bins, in three dimension 125 bins, etc. Given that each stock is placed in one and only one bin, it is doubtful whether there are enough stocks to meaningfully apply this approach beyond two dimensions. In three dimensions, the number of bins in each style dimension could be reduced say to three thus implying a total of 27 bins in the three-dimensional space. This, however, implies a less precise matching of style in each dimension. Also, suppose then that one wants to match portfolios and benchmarks in 5 dimensions. With three bins in each dimension, this requires a total of 243 bins. As the number of dimensions rises, one soon ends up with more bins than stocks. By contrast, our method works equally well in 10 style dimensions as in 1 dimension. Moreover, funds and benchmarks are exactly matched in every single dimension.

We apply our methodology to a US institutional funds data set over the period 2002–2009. The tracking error volatilities of some of our performance benchmarks are comparable with those of the CDL characteristic-matching method. This confirms the viability of our approach to benchmark construction. Also, using our new style measures, we illustrate the cross section diversity of styles across managers in the size and value-growth style dimensions and document the shift in style towards value stocks from 2002 to 2009. Finally, in contrast to the prevailing wisdom, for our sample of fund managers, we find evidence of a negative relationship between fund performance and activity (as we define it).

The remainder of the paper consists of four sections and an Appendix. Section 2 derives our fund style measures, style-matched performance benchmarks and activity measures in the context of a one-dimensional style space. Section 3 extends our approach to a multidimensional style space. Section 4 illustrates our methods using a data set of US institutional funds managers. Section 5 concludes the paper. Some derivations from Section 2 are collected in an Appendix.

2. One-dimensional performance benchmarks

2.1. A least-squares performance benchmark

Let $\tilde{w}$ denote a benchmark portfolio formed by taking linear combinations of two reference portfolios m and m^* in a given style dimension. For example, in the size dimension, m could be the Russell 3000 portfolio and m^* could be the Equal-Weighted Russell 3000 portfolio.

$$\tilde{w}_n = \lambda \times m_n + (1 - \lambda) \times m_n^* \;\; \text{for } n = 1, \ldots, N, \tag{1}$$

where $\tilde{w}_n$ denotes the value share of stock n in the benchmark portfolio, while m_n and m_n^* denote the shares of asset n in the reference style portfolios m and m^*.

By construction, $\sum_{n=1}^{N} \tilde{w}_n = \sum_{n=1}^{N} m_n = \sum_{n=1}^{N} m_n^* = 1$, for all n, where N is the number of assets in the comparison. For example, if the comparison is made over the stocks included in the Russell 3000 index, then $N = 3000$.

Now let w_n denote the value share of stock n in a portfolio of interest (e.g. that of a particular fund manager). Again, by construction $\sum_{n=1}^{N} w_n = 1$. The sum of squared deviations (SSD) between the portfolio w and its benchmark $\tilde{w}$ can be written as follows:

$$SSD = \sum_{n=1}^{N} (w_n - \tilde{w}_n)^2 = \sum_{n=1}^{N} (w_n - \lambda m_n - (1-\lambda) m_n^*)^2.$$

Choosing λ to minimize SSD yields the following first order condition:

$$\sum_{n=1}^{N} [(m_n - ew_n)(w_n - \hat{\lambda} m_n - (1-\hat{\lambda}) m_n^*] = 0. \tag{2}$$

Solving (2) we obtain the least-squares scalar:

$$\hat{\lambda}(w) = \frac{\sum_{n=1}^{N} [(m_n - m_n^*)(w_n - m_n^*)]}{\sum_{n=1}^{N} (m_n - m_n^*)^2}. \tag{3}$$

Substituting $\hat{\lambda}(w)$ into (1) yields the benchmark portfolio $\hat{w}$ that in a least-squares sense most closely approximates our portfolio of interest w.

$$\begin{aligned} \hat{w}_n &= \hat{\lambda}(w) \times m_n + [1 - \hat{\lambda}(w)] \times m_n^* \\ &= \left\{ \frac{(m_n - m_n^*) \sum_{i=1}^{N} [(m_i - m_i^*)(w_i - m_i^*)]}{\sum_{i=1}^{N} (m_i - m_i^*)^2} \right\} + m_n^*, \end{aligned} \tag{4}$$

for $n = 1, \ldots, N$. Or in matrix notation,

$$\hat{w} = \hat{\lambda} m + (1-\hat{\lambda}) m^* = \left[\frac{(m - m^*)^T (w - m^*)}{(m - m^*)^T (m - m^*)} \right] (m - m^*) + m^*.$$

This least-squares approach could easily be extended to include more than two reference portfolios in a given dimension. One reason for not doing this is that it complicates the construction of style profiles (see Section 2.2).

The performance benchmark for w is equal to the return on its benchmark portfolio, denoted here by $R(\hat{w})$. By construction:

$$R(\hat{w}) = \hat{\lambda}(w) \times R(m) + [1 - \hat{\lambda}(w)]R(m^*), \tag{5}$$

where $R(m)$ and $R(m^*)$ denote the returns on the m and m^* reference portfolios during the period of interest. If m and m^* are existing indices, such as the Russell 3000 and Russell 3000 Equal-Weighted, then $R(m)$ and $R(m^*)$ can be taken directly from the index provider (in this case Russell). Likewise, $\hat{\lambda}$ can be easily calculated from (3). Given that tracking funds are often defined on these reference portfolios, the transaction costs incurred from holding the benchmark portfolio $\hat{w}$ are generally relatively low.

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 of returns 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 cross section structure it can be applied over much shorter time intervals and linked over time to better capture time varying drifts in factor exposures.

We conclude this subsection by briefly considering two extensions on our basic method. First, when $\hat{\lambda} > 1$, this implies that the benchmark portfolio $\hat{w}$ is constructed by shorting the reference portfolio m, while when $\hat{\lambda}<0$, $\hat{w}$ is constructed by shorting the reference portfolio m^*. If a fund manager is not allowed to take short positions, it may be desirable to restrict the range of $\hat{\lambda}$ to the [0,1] interval. In this case, $\hat{\lambda}$ is replaced by $\hat{\hat{\lambda}}$, defined below, when calculating $\hat{w}$.

$$\hat{\hat{\lambda}} = \min\left[\max(0, \hat{\lambda}), 1\right]$$

As long as the reference portfolios m and m^* do not contain any short positions, restricting $\hat{\lambda}$ in this way will also ensure that $\hat{w}_n \geq 0$ for all n. When shorting is a problem, it may be desirable also to choose reference portfolios m and m^* that are far apart in the style space to ensure that for the vast majority of fund manager portfolios w we have that $0 \leq \hat{\lambda}(w) \leq 1$.

Second, when the importance of deviations between w_n and $\tilde{w}_n$ varies depending on the characteristics of each stock, it may be preferable to minimize the *weighted* sum of squared deviations (WSSD) between w and $\tilde{w}$. For example, supposing deviations between w_n and $\tilde{w}_n$ of a given magnitude have a bigger impact on overall performance when n is a small-cap stock, then small-cap stocks should be given more weight in the minimization problem. An example of such a WSSD formula is the following:

$$WSSD = \sum_{n=1}^{N}\left[\left(\frac{1 - mcw_n}{N - 1}\right)(w_n - \tilde{w}_n)^2\right].$$

2.2. The least-squares scalar as a style profile

Portfolios can be classified by style in a number of dimensions (e.g. size, growth and momentum). 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 a ranking of portfolios according to that particular style.

A natural style profile in the size dimension is the following:

$$P_{S^*}(w) = \sum_{n=1}^{N}(w_n \times mcw_n), \tag{6}$$

where mcw_n denotes the market-cap share of stock n (i.e. $\sum_{n=1}^{N} mcw_n = 1$), and w_n its value share in portfolio w as noted above. In the value-growth dimension, one possible style profile is the following:

$$P_{G^*}(w) = \prod_{n=1}^{N}(p_n/b_n)^{w_n}, \tag{7}$$

where p_n and b_n denote stock n's price and book value per share, respectively.[3]

The least-squares scalar $\hat{\lambda}$ from (3) can also be interpreted as a style profile. This style profile, denoted by $\hat{P}(w)$, is defined as follows:

$$\hat{P}(w) = \hat{\lambda} = \frac{\sum_{n=1}^{N}[(m_n - m_n^*)(w_n - m_n^*)]}{\sum_{n=1}^{N}(m_n - m_n^*)^2}. \tag{8}$$

$\hat{P}(w)$ measures the style of a portfolio w relative to the reference portfolios m and m^*. From inspection of (8), it can be seen that $\hat{P}(m^*) = 0$ and $\hat{P}(m) = 1$. Also a portfolio w and its least-squares benchmark $\hat{w}$ are exactly style matched (i.e. $\hat{P}(w) = \hat{P}(\hat{w})$).[4]

[3] One attractive feature of this formula is that its reciprocal is the market-cap-weighted geometric mean of the book-to-price ratios. 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.

[4] This exact style match no longer holds when $\hat{\lambda} \notin [0, 1]$ and the no-shorting constraint is imposed (i.e. $\hat{\lambda}$ is replaced by $\hat{\hat{\lambda}}$ when calculating $\hat{w}$ in (4)).

This result is derived from the definition of $\hat{w}$ in (4) as follows:

$$\begin{aligned}\hat{w} &= \hat{\lambda}m + (1-\hat{\lambda})m^* \\ &\Rightarrow \hat{w} - m^* = \hat{\lambda}(m - m^*) \\ &\Rightarrow (m-m^*)^T(\hat{w}-m^*) = \hat{\lambda}(m-m^*)^T(m-m^*) \\ &\Rightarrow \hat{\lambda} = \frac{(m-m^*)^T(\hat{w}-m^*)}{(m-m^*)^T(m-m^*)} \equiv \hat{P}(\hat{w}).\end{aligned} \tag{9}$$

Finally, from (8), we likewise know that $\hat{P}(w) = \hat{\lambda}$

2.3. *Choosing the reference profiles* m *and* m*

There remains the question of how the reference portfolios m and m^* should be chosen in a given style dimension. In our empirical comparisons in Section 4, in the size dimension, we set $m_n = mcw_n$ and $m_n^* = ew_n$, where *mcw* again denotes market-cap weighting and *ew* denotes equal weighting (i.e. $ew_n = 1/N$). The portfolio $m^* = ew$ here plays the role of a style neutral portfolio that can be used as one of the reference portfolios in each style dimension.[5] In this case, it can be shown that our size profile $\hat{P}_S(w)$ from (8) is a positive linear function of $P_{S^*}(w)$ (see the Appendix) and hence achieves the same ordinal ranking of portfolios as $P_{S^*}(w)$. In the value-growth dimension, we set $m_n = g\breve{w}_n$ and $m_n^* = ew_n$, where

$$g\breve{w}_n = \frac{\ln(p_n/b_n)}{\sum_{m=1}^N \ln(p_m/b_m)}. \tag{10}$$

The resulting growth profile $\hat{P}_G(w)$ is a positive monotonic function of $P_{G^*}(w)$ and hence achieves the same ordinal ranking as $P_{G^*}(w)$ (again see the Appendix).

An alternative approach, which we do not pursue further here, constructs m and m^* from the rank ordering of assets in that particular style dimension. For example, in the size dimension, suppose $r(n)$ denotes the market-cap rank of stock n. If stock n is the one with the smallest market cap, then $r(n) = 1$. If stock n is the one with the largest market cap, then $r(n) = N$. Each element of the m and m^* vectors is then calculated as follows:

[5] Alternatively, *mcw* could be treated as a style neutral portfolio. In this case, we would set $m_n^* = mcw_n^*$ in each style dimension. In the size dimension, m would then be a large-cap portfolio such as the Russell 1000 portfolio, while in the value-growth dimension m could be the Russell 3000 Growth portfolio.

$$m_n = \frac{r(n)}{\sum_{i=1}^{N} r(i)}, \quad m_n^* = \frac{N+1-r(n)}{\sum_{i=1}^{N} r(i)}. \tag{11}$$

Similarly, in the value-growth dimension, stocks could be ranked by their price/book ratios. $r(n) = 1$ now for the stock with the smallest price/book ratio, and $r(n) = N$ for the stock with the largest price/book ratio. m and m^* are then constructed again using (11). One attraction of this rank-based approach is that it prevents a small number of stocks exerting a disproportionate influence. For example, in the size dimension, when it is market-cap weighted, m may be dominated by a few large-cap stocks.

2.4. 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 per cent of the total market cap is in stock 1, 30 per cent in stock 2, 20 per cent in stock 3 and 10 per cent in stock 4. The holdings of nine hypothetical managers are shown in Table 1. For example, Manager 1 holds only stock 1, while Manager 2 holds only stock 2, etc.

Using a market-cap (*mcw*) portfolio as our reference m portfolio and an equal-weighted (*ew*) portfolio as our reference m^* portfolio, (8) generates the size profiles $\hat{P}_S(w)$ shown in Table 1. Now from (4), $\hat{\lambda}(w^1) = \hat{P}_S(w^1) = 3$ 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. $\hat{w}_n^1 = 3mcw_n - 2ew_n$). This yields the holdings $\hat{w}^1 = (0.7, 0.4, 0.1, -0.2)$. Our absolute size measure $P_{S^*}(w)$ defined in (6) can be used to verify that w^1 and $\hat{w}^1$ are the same size. $P_{S^*}(w^1) = 1 \times 0.4 = 0.4$ 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. $\hat{w}_n^4 = 4ew_n - 3mcw_n$), and hence, $\hat{w}^4 = (-0.2, 0.1, 0.4, 0.7)$. Now $P_{S^*}(w^4) = 1 \times 0.1 = 0.1$ which is the same as $P_{S^*}(\hat{w}^4) = -0.2 \times 0.4 + 0.1 \times 0.3 + 0.4 \times 0.2 + 0.7 \times 0.2 = 0.1$.

The exact size match between w^i and $\hat{w}^i$ can be demonstrated for the other managers in Table 1 in an analogous manner.

If in addition we know the return on the reference market-cap-weighted and equal-weighted portfolios, then using (5) we can compute the tailored performance benchmark for each manager. Here, we assume that the return on the market-cap-weighted portfolio is 2 per cent (i.e. $R(mcw) = 2$), while the return on the equal-weighted portfolio is 3 per cent (i.e. $R(ew) = 3$). Substituting these values into (5), along with the estimated size profile

Table 1
Simple examples consisting for four stocks and nine fund managers

	mcw	m	m^*	w^1	w^2	w^3	w^4	w^5	w^6	w^7	w^8	w^9
Stock 1	0.4	0.4	0.25	1	0	0	0	0.5	0.33	0.33	0	0.4
Stock 2	0.3	0.3	0.25	0	1	0	0	0	0.33	0.33	0.6	0
Stock 3	0.2	0.2	0.25	0	0	1	0	0	0	0.33	0.4	0.6
Stock 4	0.1	0.1	0.25	0	0	0	1	0.5	0.33	0	0	0
$\hat{\lambda} = \hat{P}_S(w)$		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
$R(\hat{w})$ in %		2	3	0	2	4	6	3	2.67	2	2.8	2.4

In these examples, there are four stocks. *mcw* denotes the market-cap weights of the stocks. Stock 1 has 40 per cent of the market cap, stock 2 has 30 per cent, stock 3 has 20 per cent and stock 4 the remaining 10 per cent. The m reference portfolio is market-cap weighted, while the m^* 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 (8). P_{S^*} is the size profile defined in (6). It is shown in the Appendix that $\hat{P}_S(w)$ is a positive linear function of $P_{S^*}(w)$. The benchmark portfolios $\hat{w}^i$ for each portfolio w^i are obtained by inserting the size profiles $\hat{\lambda}(w^i) = \hat{P}_S(w^i)$ into (4). Finally, $R(\hat{w})$ is the tailored performance benchmark for each manager. In calculating $R(\hat{w})$, we assume that the return on the m (i.e. market-cap weighted) portfolio $R(m)$ is 2 per cent while the return on the reference m^* (i.e. equal-weighted) portfolio $R(m^*)$ is 3 per cent. Given these reference returns and using the formula in (5), we obtain the performance benchmarks shown in the final row of the table. For example, $R(\hat{w}^1) = 0$ implies a benchmark return for Manager 1 of 0 per cent over the period of interest, while $R(\hat{w}^2) = 2$ implies a benchmark return for Manager 2 of 2 per cent.

$\hat{P}_S(w) = \hat{\lambda}$ of each manager, yields performance benchmarks that range from 0 per cent for Manager 1 up to 6 per cent for Manager 4.

2.5. Distance as a measure of activity

The Euclidean distance between w and its characteristic-matched benchmark $\hat{w}$ is measured as follows:

$$\hat{D} = \sqrt{\sum_{n=1}^{N} (w_n - \hat{w}_n)^2}. \tag{12}$$

We interpret $\hat{D}$ as a measure of the activity of a portfolio in a particular style dimension (where again by ‘activity’ we mean departures from passive tracking of a benchmark rather than turnover). An active portfolio can be distinguished by its deviation from its passive characteristic-matched style benchmark. $\hat{D}$ is in spirit similar 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 multidimensional style space. Second, the CP activity measure optimizes using mean absolute deviation, while we use least squares.

A variant on this index have also been used previously by Kacperczyk *et al.* (2005), henceforth KSZ, and Brands *et al.* (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 style-matched benchmark portfolios.

In Section 4.7, we explore the relationship between activity as we have defined it and performance.

3. Multidimensional performance 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 multidimensional characteristic-matched performance benchmarks can be constructed that simultaneously match a portfolio's style in each of K dimensions, indexed by $k = 1, \ldots, K$.

Suppose now that $m_k \neq m_k^*$. A linear combination $\tilde{w}$ of the $2K$ reference portfolios m_k and m_k^* for $k = 1, \ldots, K$ can be written as follows:

$$\tilde{w} = \lambda_1 m_1 + \lambda_1^* m_1^* + \lambda_2 m_2 + \lambda_2^* m_2^* + \cdots + \lambda_K m_K + (1 - \lambda_1 - \lambda_1^* - \lambda_2 - \lambda_2^* - \cdots - \lambda_K) m_K^*. \tag{13}$$

(13) can be rewritten in matrix notation as follows:

$$\tilde{w} - m_K = M\lambda, \tag{14}$$

where

$$M = (m_1 - m_K^* \; m_1^* - m_K^* \; m_2 - m_K^* \; m_2^* - m_K^* \cdots m_K - m_K^*),$$

and λ is now a $(2K-1) \times 1$ vector.

We can now write the sum of squared deviations between elements of w and $\tilde{w}$ as follows:

$$(w - \tilde{w})^T(w - \tilde{w}) = (w - m_K^*)^T(w - m_K^*) - 2\lambda^T M^T(w - m_K^*) + \lambda^T M^T M\lambda.$$

Least-squares minimization over λ yields the following first order condition:

$$M^T(w - m_K^*) = (M^T M)\hat{\lambda}. \tag{15}$$

Rearranging, we obtain that

$$\hat{\lambda} = (M^T M)^{-1} M^T(w - m_K^*). \tag{16}$$

Substituting (16) into (13) we obtain the benchmark portfolio $\hat{w}$:

$$\hat{w} = M\hat{\lambda} + m_K^* = M(M^T M)^{-1} M^T w + [I_{2K-1} - M(M^T M)^{-1} M^T] m_K^*. \tag{17}$$

It can be shown that w and $\hat{w}$ are styled matched in each dimension. It follows from (17) that

$$M^T(\hat{w} - m_K^*) = (M^T M)\hat{\lambda}. \tag{18}$$

Comparing the first order condition in (15) with (18) we now obtain that

$$M^T w = M^T \hat{w}.$$

Focusing on row i of $M^T w$ and $M^T \hat{w}$, we have that

$$(m_i - m_K^*)^T w = (m_i - m_K^*)^T \hat{w},$$

which on rearrangement becomes

$$(w - \hat{w})^T m_i = (w - \hat{w})^T m_K^*.$$

This result holds for each reference vector m_1, m_1^*, m_2, m_2^*, etc. It therefore follows that

$$(w - \hat{w})^T m_i = (w - \hat{w})^T m_i^*,$$

and hence that

$$(m_i - m_i^*)^T w = (m_i - m_i^*)^T \hat{w}.$$

Now subtracting $(m_i - m_i^*)^T m_i^*$ from both sides and dividing through by $(m_i - m_i^*)^T(m_i - m_i^*)$ we obtain that

$$\frac{(m_i - m_i^*)^T(w - m_i^*)}{(m_i - m_i^*)^T(m_i - m_i^*)} = \frac{(m_i - m_i^*)^T(\hat{w} - m_i^*)}{(m_i - m_i^*)^T(m_i - m_i^*)},$$

which reduces to

$$\hat{P}_i(w) = \hat{P}_i(\hat{w}).$$

If desired, we can set $m_k^* = ew$ for $k = 1, \ldots, K$, where the *ew* portfolio is viewed as style neutral in the sense that it treats all stocks symmetrically. In this case, we have only $K + 1$ rather than $2K$ reference portfolios. Hence

$$\tilde{w} = \lambda_1 m_1 + \lambda_2 m_2 + \cdots + \lambda_K m_K + (1 - \lambda_1 - \lambda_2 - \cdots - \lambda_K)ew. \tag{19}$$

In matrix notation:

$$\tilde{w} - ew = M\lambda, \tag{20}$$

where

$$M = (m_1 - ew\ m_2 - ew \cdots m_K - ew),$$

and λ is a $K \times 1$ vector. Replacing m_K^* with *ew* in each equation, the derivation is otherwise identical to the more general case considered above.

4. An application to institutional fund managers

4.1. 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. In any given quarter, the number of managers present is rather less than this. The highest number (i.e. 464) is observed in 2008Q4. The accounts in our data set are representative 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. The returns are gross of fees.

Our focus on institutional funds can be justified by the fact that they control more assets than mutual funds (see Christopherson *et al.*, 1998; DelGuercio

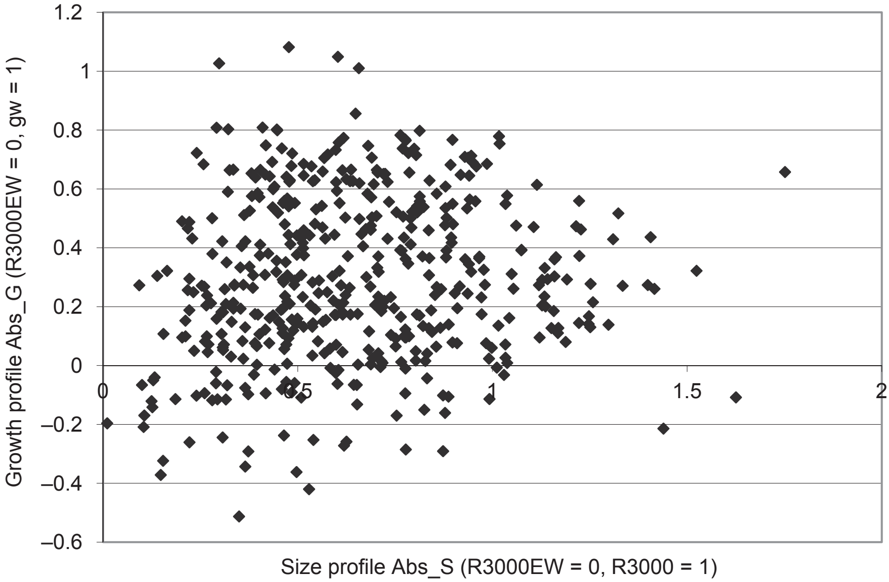

Figure 1 US fund managers: 2008Q4. This figure depicts a cross section plot of size profiles and growth profiles of all 464 fund managers in our data set in 2008Q4. Most managers have a positive growth profile (implying a tilt towards growth stocks), and size profiles between 0 and 1 (implying a size 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.

and Tkac, 2002). The data set, while not free of selection bias as larger more successful managers are more likely to be included, is based on the published Russell Mellon performance universe that is well known in the institutional investment industry. In excess of 80 per cent of institutional assets are represented in the data set (it is difficult to be more precise than this as we cannot say for certain how many managers are missing from every possible available fund). The data set therefore provides a very good coverage of the opportunity set of active US equity managers available to institutional investors.

4.2. 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 behaviour. 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 m = Russell 3000 and m^* = Russell 3000 Equal-Weighted in the size dimension, while $m = \breve{gw}$ as defined in (10) and m^* = Russell 3000 Equal-Weighted in the value-growth dimension.

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 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 $\breve{gw}$ index defined in (10).

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.

4.3. *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 w_S defined below in Section 4. 4) stays reasonably constant over the period 2002–2009, 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 w_G also defined in Section 4. 4).

The shift of fund managers towards value apparent in Figure 2b demonstrates how the maintenance of a fixed style may require the active intervention of an investor.[6] 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 behaviour changes over time. Other approaches to addressing these issues, such as risk models and equity profiles, are expensive to purchase, computationally more intensive and not necessarily as easy to interpret.

4.4. *Description of benchmarks and empirical results*

Average tracking error volatilities of characteristic-matched performance benchmarks, gross excess returns and activity measures for funds relative to these benchmarks are shown in Table 2.

The tracking error of a fund is calculated as the annualized standard deviation of quarterly excess returns versus its tailored benchmark. This measure is often used in the literature and the investment industry to judge the appropriateness of benchmarks (see, for example, CDL and CP). A lower tracking error implies a closer match between the fund and benchmark, suggesting that the benchmark is a better representative of the fund's habitat than benchmarks with higher tracking errors.

When calculating tracking error volatilities, we only consider managers for which we have at least 12 consecutive quarters of data. This reduces our sample

[6] See Ainsworth *et al.* (2008) for an in-depth discussion of style drift and its implications for investors.

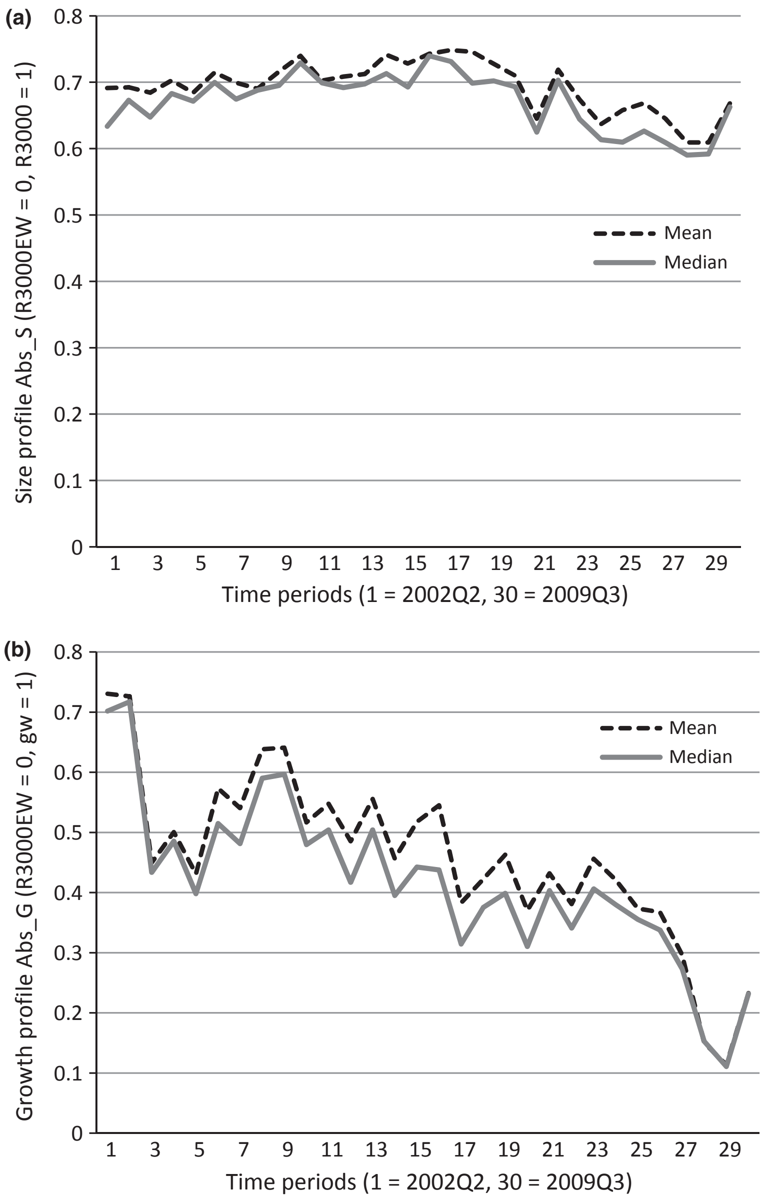

Figure 2 Panel (a) The evolution of the average fund size profile. Panel (b) The Evolution of the Average Fund Growth Profile. Panel 'a' shows how the average size profile (Abs_S) of funds in our data set was reasonably stable over the 2002Q2–2009Q3 period. Panel 'b' shows how the average growth profile (Abs_G) has a clear downward trend, implying a strong shift from growth to value stocks.

Table 2
Benchmark tracking error volatilities, excess returns and activity for funds (2002Q2–2009Q3)

	CDL	w_S	w_G	w_SG	w_G(mcw)
TE volatility					
Median	4.4690	5.5101	7.5945	4.8505	4.4663
SD	1.7807	2.1238	1.5361	1.7152	1.9764
Min	1.2717	2.1075	4.4849	2.1481	1.3451
Max	11.9272	14.6077	13.5771	11.2385	12.5606
Excess return					
Median	1.4247	0.3795	−0.2436	0.7309	0.8499
SD	2.4661	2.5644	2.6829	2.5288	2.5497
Min	−7.5056	−9.7014	−9.8622	−11.1001	−10.9438
Max	9.4969	6.4551	6.9226	7.7626	7.4485
Activity					
Median	0.1346	0.1302	0.1423	0.1302	0.1306
SD	0.0322	0.0334	0.0335	0.0333	0.0313
Min	0.0918	0.0791	0.0833	0.0789	0.0912
Max	0.3279	0.3223	0.3229	0.3222	0.3275

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 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 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 w_S performance benchmark portfolio is calculated by taking a linear combination of the Equal-Weighted Russell 3000 and Russell 3000 portfolios. The w_G portfolio is a linear combination of the Equal-Weighted Russell 3000 and $g\breve{w}$ portfolios. The w_SG portfolio is a linear combination of the Equal-Weighted Russell 3000, Russell 3000 and $g\breve{w}$ portfolios. The w_G(mcw) portfolio is a linear combination of the Russell 3000 Value and Russell 3000 Growth portfolios.

of managers in Table 2 to 275. The choice of the number of consecutive quarters required for inclusion is somewhat subjective. CDL, for example, require 16 consecutive quarters. We prefer 12 as the gains in sample size (275 instead of 164 funds) in our opinion outweigh 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.

The first set of results in Table 2 replicates the value-weight conditional sort (i.e. quarterly size, within-size, BM) method used by CDL. This is their preferred method as 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.[7] Hence, we use the CDL value-weight conditional sort method 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 2. Our four benchmarks are described below.

1. w_S: one-dimensional size benchmark: m = Russell 3000, m^* = Russell 3000 Equal-Weighted
2. w_G: one-dimensional growth benchmark: $m = \breve{gw}$ (as defined in (10)), m^* = Russell 3000 Equal-Weighted
3. w_SG: two-dimensional size-growth benchmark: m_S = Russell 3000, m^*_S = Russell 3000 Equal-Weighted, $m_G = \breve{gw}$, m^*_G = Russell 3000 Equal-Weighted
4. w_G(mcw): one-dimensional growth benchmark: m = Russell 3000 Growth, m^* = Russell 3000 Value

The last of these requires some explanation. w_G(mcw) provides a growth benchmark for broadly market-cap-weighted funds. This is because the reference portfolios Russell 3000 Growth and Russell 3000 Value, while differing in the value-growth dimension, are both essentially market-cap weighted. The difference between the w_G and w_G(mcw) portfolios can be seen by comparing their size profiles $P_S(w_G)$ and $P_S(w_G(mcw))$. w_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 w_G(mcw) has a size profile similar to that of the Russell 3000 index (i.e. close to 1). In other words, w_G provides a benchmark for growth/value tilts relative to an equal-weighted portfolio, while w_G(mcw) provides a benchmark for growth/value tilts relative to a market-cap-weighted portfolio. Some implications of this distinction are explored below.

4.5. Gross excess returns of funds

To interpret the pattern of gross excess returns in Table 2, it is first useful to consider the annualized average total percentage returns on the reference

[7] 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, and hence to improve comparability likewise do not consider CDL's composite value-growth measures either.

Russell indices (including dividends and other payments). 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.[8] The difference in the performance of the Russell 3000 and Russell 3000 Equal-Weighted indices is quite striking. By implication, a performance benchmark for small-cap managers needs to be much higher than for large-cap managers over this sample period. While the gap here is perhaps bigger than usual, a large literature has confirmed that small-cap portfolios tend to outperform large-cap portfolios (see, for example, Chan *et al.*, 2009a).

The average fund in our data set has an w_S size profile of about 0.68 (see Figure 2a), which lies in between that of the Russell 3000 Equal-Weighted and Russell 3000 indices (which by construction have w_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, 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 1.31 per cent per year (not shown in Table 2) but only 0.38 per cent relative to w_S.

The poor performance of funds relative to w_G (as evidenced by an average excess return of −0.24 per cent per year in Table 2) can be explained by the $g\breve{w}$ 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 0.68, there is a clear size mismatch between funds and their growth-adjusted w_G benchmarks (even though their style is matched in the value-growth dimension). Given that small stocks significantly outperformed large stocks over our sample period, it is therefore not surprising that funds on average underperformed their w_G benchmarks.

Average fund performance improves dramatically from −0.24 per cent to +0.85 per cent when the benchmark is switched from w_G to w_G(mcw). This is because the w_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 w_G(mcw) 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 w_G(mcw) portfolios. As a result, funds on average outperform their w_G(mcw) 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

[8] The annual rates of return here differ slightly from those on the Russell website as we have excluded a few stocks for which we could not obtain book values which are required to calculate the w_G and w_SG benchmarks.

meaningful results are obtained by matching style simultaneously on size and value-growth. Such matching is achieved by our w_SG benchmark. According to w_SG, the average gross excess return of funds in our data set is +0.73 per cent per year.

4.6. Tracking error volatility of benchmarks

Median tracking error volatilities are presented in Table 2. The best performer is w_G(mcw), followed in order by CDL, w_SG, w_S, and lastly w_G.[9]

The relatively high tracking error of w_G is probably caused by it adjusting more over time than the fund manager portfolios themselves (cap mismatch may be the biggest driver here). This is because the price/book ratio of each stock is generally more volatile than its market-cap share. The fact that both w_G(mcw) and w_SG have lower tracking errors than w_S suggests that value-growth is a major driver of shifts in funds' approaches to stock picking.

The results in Table 2 demonstrate that at least some of our methods are competitive in terms of tracking error volatility with CDL's preferred method.

In higher dimensions, our methods should perform even better relative to existing characteristic-matching methods (including CDL's preferred method), as the latter will then be forced to use coarser categories in each style dimension to prevent the total number of bins becoming too large. 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.[10]

[9] Exactly the same median rankings of methods is obtained when the comparison is restricted to funds present for at least 16 consecutive quarters.

[10] Brands *et al.* (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.

4.7. 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 (12).[11] Activity quintiles and their corresponding average gross excess returns calculated using our four characteristic-matched benchmarks – w_S, w_G and w_SG, w_G(mcw) – and the Russell 3000 and CDL portfolios as benchmarks are shown in Table 3.

In Table 3, 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 3, where the activity of each fund is measured relative to its own characteristic-matched benchmark using the formula in (12). 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 gross excess returns are then averaged across all quarters. The *t*-statistics for the excess returns in Table 3 are calculated using the heteroscedasticity and autocorrelation consistent (HAC) estimator of Newey and West (1987).

It should be noted that Table 3 is calculated using all the 1183 funds in the data set, while Table 2 uses only the 275 funds that were present for at least 12 consecutive quarters. We are able to include all funds in Table 3 as we are not computing tracking error volatilities. One implication of this is that the excess return and activity results in Tables 2 and 3 are not directly comparable. Also, the quarterly excess returns in Table 3 have not been annualized.

The results in Table 3 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 per cent significance level. Nevertheless, the general pattern in Table 3 is reasonably clear.

[11] We could in principle also use our approach to investigate the link between tracking error volatility and performance.

Table 3
Activity versus performance

2002Q2–2009Q3	Activity quintiles: Lowest	2nd	3rd	4th	Highest	Low–high
w_S						
Activity	0.0988	0.1202	0.1360	0.1536	0.1951	0.2929
Excess return (%)	−0.0221	−0.0385	−0.1373	−0.2504	−0.3150	
t-Stat	−0.0874	−0.1889	−0.7599	−1.1787	−1.3613	
w_G						
Activity	0.1076	0.1313	0.1476	0.1650	0.2042	0.1925
Excess return (%)	−0.2839	−0.3633	−0.4916	−0.5085	−0.4764	
t-Stat	−0.6035	−0.6967	−0.9472	−1.0504	−1.0673	
w_SG						
Activity	0.0988	0.1200	0.1358	0.1534	0.1948	0.0636
Excess return (%)	0.0530	0.0643	−0.0066	−0.0321	−0.0106	
t-Stat	0.2483	0.4248	−0.0517	−0.2396	−0.0704	
w_G(mcw)						
Activity	0.1032	0.1205	0.1355	0.1531	0.1971	0.1935
Excess return (%)	0.2792	0.2862	0.1938	0.0939	0.0857	
t-Stat	1.2840	1.5882	1.2586	0.5164	0.4911	
R3000						
Activity	0.1051	0.1242	0.1390	0.1565	0.1978	0.2528
Excess return (%)	0.4146	0.4343	0.2692	0.1264	0.1618	
t-Stat	1.9093	2.7231	1.7375	0.7682	0.8672	
CDL						
Activity	0.0635	0.0825	0.0976	0.1153	0.1538	0.3512
Excess return (%)	0.5680	0.3838	0.2877	0.2495	0.2168	
t-Stat	1.9593	1.8131	1.6446	1.2772	1.1660	
w_S						
Activity	0.0994	0.1205	0.1362	0.1532	0.1957	0.5190
Excess return (%)	0.1039	0.0267	−0.1022	−0.2171	−0.4152	
t-Stat	0.3296	0.1278	−0.5035	−1.2241	−2.3989	
w_G						
Activity	0.1091	0.1321	0.1486	0.1655	0.2055	0.5177
Excess return (%)	−0.2256	−0.4248	−0.5803	−0.6476	−0.7433	
t-Stat	−0.3707	−0.6906	−0.9983	−1.2043	−1.5210	
w_SG						
Activity	0.0994	0.1203	0.1360	0.1530	0.1954	0.3419
Excess return (%)	0.1863	0.0618	0.0602	−0.0310	−0.1555	
t-Stat	0.6746	0.2995	0.3779	−0.2249	−1.1044	
w_G(mcw)						
Activity	0.1034	0.1207	0.1353	0.1529	0.1987	0.1811
Excess return (%)	0.4257	0.4307	0.3536	0.1750	0.2447	
t-Stat	3.3449	4.0698	2.9784	1.3468	1.5314	
R3000						
Activity	0.1050	0.1244	0.1390	0.1562	0.1985	0.4115
Excess return (%)	0.4660	0.5047	0.2710	0.0751	0.0544	
t-Stat	3.4080	4.3653	2.3660	0.5083	0.2558	

Table 3 (continued)

	Activity quintiles					
2002Q2–2009Q3	Lowest	2nd	3rd	4th	Highest	Low–high
CDL						
Activity	0.0633	0.0827	0.0986	0.1168	0.1555	0.4832
Excess return (%)	0.6816	0.4661	0.2803	0.3121	0.1984	
t-Stat	3.3418	2.6129	1.8585	2.0779	1.1293	

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 2. In addition, activity measures and excess returns by quintile are also calculated using the Russell 3000 and CDL portfolios as benchmarks.

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 2002Q2–2009Q3, while CP's covers the period 1990–2003. 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. This lack of overlap could be quite important given the different environments in which institutional funds and mutual funds operate (Christopherson *et al*., 1998). 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 3, 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 or CDL benchmarks also does not change the general thrust of our results.

This suggests that further work is necessary to validate the strong claims in CP over different time periods, asset classes and fund types (e.g. mutual funds versus institutional funds) and may give pause to investors seeking to use the difference in holdings of a fund from its simple benchmark as a way ot predict future outperformance.

5. Summary and conclusions

Characteristic-matched performance benchmarks obtained from portfolio holdings data are typically constructed using a bottom-up approach that 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. These reference portfolios can often be taken off-the-shelf.

The fact that our approach matches funds and benchmarks at the macrolevel means that it is capturing a different aspect of the style-matching problem than traditional characteristic-matching methods that match at the microlevel. In this sense, it should be viewed as a complement to existing methods.

Furthermore, our approach has some distinct advantages. First, it also simultaneously generates measures of a fund's style and activity. Second, its benchmarks are potentially investable at lower cost, particularly when the reference indices have tracking funds. Third, its style matches in each dimension are exact. Fourth, it avoids the curse of dimensionality. By contrast, standard characteristic-matching methods in higher dimensions soon run into the problem of having too many bins relative to the number of stocks.

Applying our new style-profiling benchmarks and measures of style and activity to US institutional funds data over the period 2002–2009, we observe three main results. First, we document the drift in average fund style over time towards value and find that the average fund holds a portfolio smaller in size than the Russell 3000 portfolio. Second, the tracking error volatilities of at least some of our style-profiling benchmarks even when applied in only one or two style dimensions are comparable with those of standard characteristic-matching methods. In higher dimensions, the relative performance of our style-profiling benchmarks should be even better. Finally, we observe a negative relationship between activity and performance.

References

Ainsworth, A. B., K. Fong, and D. R. Gallagher, 2008, Style drift and portfolio management for active Australian equity funds, *Australian Journal of Management* 32 (3), 387–418.

Annaert, J. and G. Van Campenhout, 2007, Time variation in mutual fund style exposures, Review of Finance 11(4), 633–661.

Bailey, J. V., 1992, Are manager universes acceptable performance benchmarks?, *The Journal of Portfolio Management* 18(3), 9–13.

Brands, S., S. J. Brown, and D. R. Gallagher, 2005, Portfolio concentration and investment manager performance, *International Review of Finance* 5(3–4), 149–174.

Brown, S. J., and W. N. Goetzmann, 1997, Mutual fund styles, *Journal of Financial Economics* 43(3), 373–399.

Carhart, M. M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52 (1), 57–82.

Chan, L. K. C., H. L. Chen, and J. Lakonishok, 2002, On mutual fund investment styles, *Review of Financial Studies* 15(5), 1407–1437.

Chan, H., R. W. Faff, D. R. Gallagher, and A. Looi, 2009a, Fund size, transaction costs and performance: size matters! *Australian Journal of Management* 34(1), 73–96.

Chan, L. K. C., S. G. Dimmock, and J. Lakonishok, 2009b, Benchmarking money manager performance: Issues and evidence, *Review of Financial Studies* 22(11), 4553–4599.

Christopherson, J. A., W. E. Ferson, and D. A. Glassman, 1998, Conditioning manager alphas on economic information: Another look at the persistence of performance, *Review of Financial Studies* 11(1), 111–142.

Cremers, M., and A. Petajisto, 2009, How active is your fund manager? A new measure that predicts performance, *Review of Financial Studies* 22(9), 3329–3365.

Daniel, K., M. Grinblatt, S. Titman, and R. Wermers, 1997, Measuring mutual fund performance with characteristic-based benchmarks, *Journal of Finance* 52(3), 1035–1058.

Del Guercio, D., and P. A. Tkac, 2002, The determinants of the flow of funds of managed portfolios: Mutual funds vs. pension funds, *Journal of Financial and Quantitative Analysis* 37(4), 523–557.

Faff, R. W., 2003, Creating Fama and French factors with style, *The Financial Review* 38(2), 311–322.

Fama, E. F., and K. R. French, 1992, The cross section of expected stock returns, *Journal of Finance* 47(2), 427–465.

Fama, E. F., and K. R. French, 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51(1), 55–84.

Fong, K., D. R. Gallagher, and A. D. Lee, 2008, Benchmarking benchmarks: Measuring characteristic selectivity using portfolio holdings data, *Accounting and Finance* 48(5), 761–781.

Huij, J., and M. Verbeek, 2009, On the use of multifactor models to evaluate mutual fund performance, *Financial Management* 38(1), 75–102.

Kacperczyk, M., C. Sialm, and L. Zheng, 2005, On the industry concentration of actively managed equity mutual funds, *Journal of Finance* 60(4), 1983–2011.

Kothari, S. P., and J. B. Warner, 2001, Evaluating mutual fund performance, *Journal of Finance* 56(5), 1985–2010.

Newey, W. K., and K. D. West, 1987, A simple, positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55(3), 703–708.

Russell Investments, 2008, US equity indexes: Institutional benchmark survey.

Sharpe, W. F., 1988, Determining a fund's effective asset mix, *Investment Management Review* 2(6), 59–69.

Sharpe, W. F., 1992, Asset allocation, management style and performance measurement, *The Journal of Portfolio Management* 18(2), 7–19.

Wermers, R., 2000, Mutual fund performance: An empirical decomposition into stock picking talent, style, transaction costs, and expenses, *Journal of Finance* 55(4), 1655–1695.

Appendix

Showing that $P_S(w)$ and $P_G(w)$ are monotonic functions of $P_{S^*}(w)$ and $P_{G^*}(w)$

Let *mcw* and *ew* denote, respectively, a market-cap-weighted and equal-weighted portfolio. Setting m=mcw and m^*=ew in (8), we obtain that

$$P_S(w) = \frac{\sum_{n=1}^{N}[(mcw_n - ew_n)(w_n - ew_n)]}{\sum_{n=1}^{N}(mcw_n - ew_n)^2}, \tag{21}$$

where

$$mcw_n = \frac{p_n q_n}{\sum_{m=1}^{N} p_m q_m}, ew_n = \frac{1}{N}, \text{for } n = 1, \ldots, N.$$

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 $P_S(w)$ below is not defined.

That $P_S(w)$ as defined in (21) is a monotonic (linear) function of the size profile $P_{S^*}(w)$ defined in (6) can be demonstrated as follows:

$$\begin{aligned} P_S(w) &= \frac{\sum_{n=1}^{N}[(mcw_n - ew_n)(w_n - ew_n)]}{\sum_{n=1}^{N}(mcw_n - ew_n)^2} = \frac{\sum_{n=1}^{N}(w_n mcw_n) - 1/N}{\sum_{n=1}^{N}(mcw_n)^2 - 1/N} \\ &= \frac{P_{S^*}(w) - 1/N}{\sum_{n=1}^{N}(mcw_n)^2 - 1/N}. \end{aligned} \tag{22}$$

As long as the same list of stocks is used when computing $P_S(w)$ for all portfolios, then $\sum_{n=1}^{N}(mcw_n)^2$ and N are constants as they do not depend on w. The term $\sum_{n=1}^{N}(mcw_n)^2 - 1/N$ is 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, $P_S(w)$ is an increasing linear (and hence monotonic) function of $P_{S^*}(w)$.

Setting $m = \breve{gw}$ and $m^* = ew$ in (8), where $\breve{gw}$ is the growth weighted portfolio defined in (10), and now assuming there exist at least two stocks for which $ew_n \neq \breve{gw}_n$, we obtain that

$$\begin{aligned} P_G(w) &= \frac{\sum_{n=1}^{N}[(\breve{gw}_n - ew_n)(w_n - ew_n)]}{\sum_{n=1}^{N}(\breve{gw}_n - ew_n)^2} = \frac{\sum_{n=1}^{N}(w_n \breve{gw}_n) - 1/N}{\sum_{n=1}^{N}(\breve{gw}_n^2) - 1/N} \\ &= \frac{\ln[P_{G^*}(w)]/[\sum_{n=1}^{N}\ln(p_n/b_n)] - 1/N}{\sum_{n=1}^{N}(\breve{gw}_n^2) - 1/N}. \end{aligned} \tag{23}$$

$P_{G^*}(w)$ is the growth profile defined in (7). 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 $\sum_{n=1}^{N} (g\breve{w}_n)^2$ are constants (i.e. they do not depend on w). Hence, $P_G(w)$ is an increasing monotonic function of $P_{G^*}(w)$.

The growth variant of the normalized Herfindahl–Hirschman index $\sum_{n=1}^{N} (g\breve{w}_n)^2 - 1/N$ must be greater than zero, 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 $g\breve{w}_n$ will be negative for stocks with a price-to-book value of less than one.