ELSEVIER Journal of Financial Economics 43 (1997) 373–399

JOURNAL OF Financial ECONOMICS

Mutual fund styles

Stephen J. Brown[a], William N. Goetzmann*[,b]

[a]*Leonard Stern School of Business, New York University, New York, NY 10012, USA*
[b]*Yale School of Management, Yale University, New Haven, CT 06520, USA*

(Received September 1995; final version received June 1996)

Abstract

Mutual funds are typically grouped by their investment objectives or the 'style' of their managers. We propose a new empirical to the determination of manager 'style.' This approach is simple to apply, yet it captures nonlinear patterns of returns that result from virtually all active portfolio management styles. Our classifications are superior to common industry classifications in predicting cross-sectional future performance, as well as past performance, and they also outperform classifications based on risk measures and analogue portfolios. Interestingly, 'growth' funds typically break down into several categories that differ in composition and strategy.

Key words: Mutual funds; Management style; Style analysis
JEL classification: G20; G23; G11

1. Introduction

Investment objectives and style classifications are widely used in the financial industry to characterize differences between money managers. Mutual funds, for

*Corresponding author.

The authors thank Michael Barclay, Ken French, Mark Grinblatt, Toshiyuku Otsuki, and Matthew Spiegel for helpful comments. We also thank participants in the 1995 WFA session on investment styles, the 1995 Conference on Finance and Accounting at the University of New South Wales, the 1995 Conference on Finance and Accounting, workshops at Berkeley, The City University of Hong Kong, University of California at Irvine, International University of Japan, University of Melbourne, Rutgers, Stanford, Princeton, Virginia Polytechnic, Washington University in St. Louis, the University of Washington, and the University of Wisconsin at Milwaukee, for their suggestions and comments. We also thank Morningstar, Inc. and Ibbotson Associates for providing data for analysis. All errors are the sole responsibility of the authors.

PII S0304-405X(96)00898-7

instance, are typically grouped according to the type of securities in which they invest and the 'style' of their managers. Equity funds range from 'aggressive growth' funds holding low-dividend, high-growth stocks to 'income' funds seeking high-yield equities. Such fund classifications are ubiquitous, but do they actually tell us anything about the strategies of investment managers? Do they help explain differences in future returns among funds or even provide useful benchmarks for evaluating relative past performance? These fundamental questions about mutual fund classifications are the motivation for this research.

The definition of standard equity mutual fund categories is generally broad enough to allow a wide range of different investment policies. The Investment Company Institute uses a very general description of the largest investment category:

> Growth Funds invest in the common stock of well established companies. Their primary aim is to produce an increase in the value of their investments (capital gains) rather than a flow of dividends. (Investment Company Institute, 1991, p. 12)

This definition makes it obvious that the typical growth fund manager has great latitude in the types of stocks to hold, the timing of purchases and sales, the level of fund diversification, the industry concentration of the portfolio, and a host of other factors that go into determining the returns to client investments. Given this broad latitude, it is not surprising to find widely divergent behavior among funds pursuing the same objective. As a result, existing classifications do a poor job of forecasting differences in future performance.

Moreover, the financial press has identified several cases of funds apparently misclassifying themselves (e.g., Donnelley, 1992). The S.E.C. has a stated mandate to insure that the composition of a fund does not contradict its objective, if the objective is included as part of its name. Such governmental concerns are not unfounded. Recent papers by Witkowski (1994) and Kim, Shukla, and Tomas (1995) find that the movement of many mutual funds is better explained by the performance of a style index other than their own. In this paper, we find some evidence suggesting that such misclassification may be intentional, in that it works to improve ex post *relative* performance measures, on average.

Because management styles are so widely used as the basis for performance measurement and compensation, there is a great need for style classifications that are objectively and empirically determined, consistent across managers, and related to the manager's strategy. Objectivity is important because of the moral hazard inherent in allowing managers to self-report their styles without objective verification. Consistency is needed for purposes of performance comparison. The desirability of such a classification scheme is clear to all participants in the industry, and industry alternatives to the existing classification procedures have already begun to evolve (see, e.g., Tierney and Winston, 1991; Christopherson, 1995). Beyond the practical need for meaningful styles, there is a fundamental question about whether any classification system (which, after all, is only a multinomial statistic) is sufficient to characterize differences in fund management.

To examine all of these issues, we develop a 'style classification' algorithm that is consistent with asset pricing models. The consistency is useful, because the multinomial statistic represents a 'coarsening' of a fully specified stochastic model of portfolio returns and it is useful to clarify where and how this coarsening takes place. Our algorithm groups funds based on the cross-sectional time series of past returns as well as on the response to exogenously specified and endogenously determined stochastic variables. Using mutual fund data from 1976 through 1994, we find that equity funds broadly fall into some familiar and not-so-familiar patterns of behavior. The familiar patterns include 'small-cap', 'growth', 'growth and income', 'income', and 'international' funds. However, as many as half of all currently classified 'growth' funds fall into different categories, according to our procedure. We also identify some unfamiliar categories that are *not* captured by the traditional objectives, including 'value' managers, 'trend-chasers', and 'glamour' managers.

Our derived classifications specified ex ante do a better job of predicting cross-sectional variation in fund returns than do traditional mutual fund classifications. In addition, we find that simple classifications capture major differences in manager behavior as manifested in the temporal pattern of returns. While these classifications provide less information about the magnitude of fund loadings on major macroeconomic factors, they provide a useful means to identify widespread, common patterns in manager behavior.

The implications of our results are broad. A return-based classification system such as ours can reduce the incentive to 'game' the styles to improve relative ex post rankings. More formal classification procedures for mutual funds can help investors better understand the future behavior of their investments, and can provide ex post or ex ante performance benchmarks. From the perspective of researchers interested in understanding investment manager behavior, categories that incorporate 'value' and 'glamour' funds may better characterize how managers behave. A by-product of the estimation procedure is the creation of a parsimonious set of robust factors composed of positive weights on existing mutual funds. These style factors typically outperform prespecified macroeconomic factors in out-of-sample tests on fund returns, and thus may have further implications for asset pricing.

The paper proceeds as follows. Section 2 provides background on some statistical and strategic issues in the paper. Section 3 describes the data and methodology. Section 4 reports the results of the empirical analysis and Section 5 concludes.

2. Statistical and strategic issues

2.1. Statistical issues: Time-varying portfolio weights

There is a long tradition of characterizing mutual funds according to parameters estimated via a linear model of returns. The technology of asset pricing was first applied by Jensen (1968) to grouping mutual funds according to their systematic risk characteristics. Connor and Korajczyk (1986), Lehmann and Modest (1987), and Grinblatt and Titman (1988, 1989) all apply linear asset pricing methods to differentiate mutual funds on the basis of systematic risk characteristics. Quantitative methods of security aggregation first appeared in the finance literature in the work of Elton and Gruber (1970), who developed a classification algorithm using a linear model of fundamental characteristics that proved useful for grouping securities and forecasting cross-sectional differences. Carleton and McGee (1970) suggest that the related techniques of switching regressions and hierarchical clustering methods can be used to aggregate financial assets.

There are reasons to expect that linear models might poorly characterize mutual fund returns. Single-factor and multiple-factor linear models are only exactly correct when portfolio weights remain fixed through time and when the systematic risk characteristics of the securities held in the portfolio remain fixed as well. While it is common to assume that securities change little, we have no basis for presuming that portfolio weights remain fixed. Indeed, active fund management clearly implies the strategic reallocation of portfolio weights across assets. While typical examples of such active allocation are market timing and portfolio insurance strategies, a buy-and-hold strategy as well as active security selection can also result in time-varying systematic risk characteristics. Recent studies of mutual fund manager behavior report unambiguous evidence of strategic changes in mutual fund portfolios. Grinblatt, Titman, and Wermers (1993) identify herding activity by mutual fund managers. Ferson and Schadt (1996) find that managers rebalance in anticipation of changing economic conditions. Brown, Harlow, and Starks (1993) find systematic changes in risk conditional on past performance. Lakonishok, Shleifer, Thaler, and Vishny (1991) find that 'window dressing' accounts for portfolio rebalancing by pension fund managers.

Active portfolio management affects performance measurement in nontrivial ways. Dybvig and Ross (1985), for instance, show how linear risk models fail to properly rank fund managers when they change their asset weights through time. Connor and Korajczyk (1991) consider how to risk-adjust for nonlinear portfolio strategies by mutual fund managers. Grinblatt and Titman (1993) avoid problems posed by nonlinearities by explicitly considering active strategies as the basis for a benchmark-free approach to performance measurement.

While such nonlinearities present problems for style identification as well, our procedure accommodates nonlinear strategies by allowing factor loadings to change on a month-by-month basis. This is crucial in light of the fact that many fund managers actively vary their exposure both to the market and to industry sectors. To the extent that groups of managers change these exposures together (i.e., they 'herd' into the market, or in and out of sectors), our procedure will group them together. Although it is a relatively simple technique, when we compare the style categories formed in the space of past returns to alternate categorization schemes formed in the space of fixed factor loadings, we find the former to be superior in explaining the out-of-sample cross-section of mutual fund returns. Our method, which relies on a low-dimensional multinomial statistic with intuitive interpretation as a style, compares favorably to the use of continuous multivariate measures such as factor loadings. We find some evidence, in the form of time-varying factor loadings, that this is due to the presence of dynamic management styles in the mutual fund universe.

2.2. *Strategic issues: Self-misclassification*

Thus far we have been concerned with the ability of existing style classifications to pick up management behavior that is dynamic and not well captured by static models of investment. Of great concern is the further issue of selecting a procedure that prevents ex post changes in style in order to improve relative historical performance. A self-reported fund objective, announced ex post, could have been chosen to minimize poor relative performance. Anecdotal evidence (cited above) from the financial press suggests that such misrepresentation occurs.

We find some empirical evidence to back up the casual observation that funds can switch to improve their relative historical rankings. Using equity mutual fund data over the period 1976 through 1992, described in further detail below, we find 237 cases in which equity mutual funds switched their fund objective. For each of these, we subtract the average objective return from the fund return in the year before the switch, using first the *old* objective and then the *new* objective. That is, the net gain for fund I is defined as $(r_{i,t} - r_{j,\text{old}}) - (r_{i,t} - r_{j,\text{new}})$, where $j_{i,\text{old}}$ is the style from which the fund switched in period $t+1$ and $j_{i,\text{new}}$ is the style to which the fund switched. Thus, the difference between these is the net gain or loss in ex post performance of the previous year. The average net gain in benchmarked returns was 0.098, or 9.8%, with a t-statistic of 5.47, assuming that all switches are independent.

While this simple test does not prove that fund managers were switching for strategic purposes during this period, the results are certainly consistent with such an interpretation. Were we to use self-reported styles for benchmarking, without checking to see whether the fund recently reclassified itself, we might be misled regarding the relative performance of the fund. This is also true if we were to base the style classification of the fund on its current portfolio holdings. 'Window-dressing' is a common end-of-period ploy of fund managers to throw out poor performers and/or change the apparent strategy of the fund. Since our procedure uses past returns, not portfolio holdings, it is not fooled by window-dressing. In most cases, even if we knew that the fund switched, mutual fund data vendors do not provide a historical record of past fund classifications, so the true benchmarked history is impossible to reconstruct.

3. Data and methodology

Morningstar, Inc. provided monthly returns of equity mutual funds for the period January 1976 though June 1995, together with a classification into fifteen categories: equity income, growth and income, growth, small-company, Europe, foreign, world, Pacific, financial sector, health sector, natural resources sector, precious metals sector, high technology sector, utilities sector, and an unaligned sector. These equity categories include funds that invest in bonds as well as stocks. The distinction between equity funds and bond funds is generally one of degree. Even all-equity funds typically hold some cash balances. Like most data sources of mutual fund returns, the Morningstar data are not free of survivorship bias, and the effect of fund attrition has an unknown effect on ex post classification. In order to address the problem of changes in fund classification, we merged the Morningstar data with the annual Weisenberger data used by Brown and Goetzmann (1995). The Weisenberger data are updated through 1992, the last volume in which Weisenberger provides a comprehensive *Panorama* section to their mutual fund annual, based on funds that were willing to report their performance results over the previous year. While not entirely free of bias, the data identify changes in fund objectives through time. In addition, we use a third source of mutual fund data that provides rich material for cross-sectional analysis: the Morningstar 'On-Disc' database. While only available since 1993, this CD-ROM program provides information on the composition of each fund as well as summary statistics about the securities in the fund. We cross-index this information with the monthly returns and the Weisenberger data to allow an analysis of our endogenously determined styles by a broad range of characteristics.

3.1. Stochastic specification

The objective of our analysis is to use past returns to determine a natural grouping of funds that has some predictive power in explaining the future cross-sectional dispersion in fund returns. Such groupings are referred to as *styles*. If there are K styles, the ex post total return in period t for any fund can be represented as:

$$R_{jt} = \alpha_{Jt} + \beta' I_t + e_{jt}\,, \tag{1}$$

where fund j belongs to style J. There are several ways of interpreting this equation. In a traditional financial economics framework, this equation refers to a multi-factor or a multi-beta model. The factor loadings on the factors I_t are given by β_{Jt}. These loadings are allowed to change through time. As Sharpe (1992) points out, if we regard the factors I_t as returns on index portfolios, the factor loadings can be thought of as equivalent portfolio weights associated with a dynamic portfolio strategy that might be associated with the style in question. In an interpretation closer to that of financial practitioners, β_{Jt} refers to a characteristic of a typical stock in the Jth style classification (size, market-to-book, price–earnings multiple, etc.) and I_t is the return to that attribute (see e.g., Lakonishok, Shleifer, and Vishny, 1994).

Regardless of how we interpret the equation, the style classifications will explain the cross-sectional dispersion of fund returns. We can write the equation as

$$R_{jt} = \mu_{Jt} + \varepsilon_{jt}\,, \tag{2}$$

where μ_{Jt} is the expected return for style J conditional on the factor realization I_t. If the idiosyncratic return component ε_{jt} has a zero mean ex ante and is uncorrelated across securities, the classification into styles will suffice to explain the cross-sectional dispersion of fund returns to the extent that μ_{Jt} differs across styles.

The task of assigning funds to style categories can be thought of as a problem in endogenously defining regimes (see, e.g., Quandt, 1959, 1960). In this way, it bears a 'family' resemblance to a switching regression, although unlike the switching regression, an exact solution to the style classification problem is only obtained through exhaustive combinatorics. The approach we use finds a local optimum via the minimization of a 'within-group' sum-of-squares criterion over a specific time period, $t = 1, \ldots, T$. The inputs to the procedure are a T-by-N matrix of monthly returns for a set of N mutual funds. We group the N funds together into K styles by minimizing the within-style mean returns for each period, $t = 1, \ldots, T$. Thus, we are jointly estimating the time series of mean returns for the styles $J = 1, \ldots, T$ (μ_{Jt}) for $t = 1, \ldots, T$, and the membership to each style. The benefit of the resulting classification is that groups could result from either fixed portfolio strategies, such as similar asset compositions, or from dynamic portfolio strategies, such as portfolio insurance rebalancing.

The classification procedure assumes that we know the number of styles. Conditional upon restrictions on the exact number of groups, Eq. (2) is perfectly well-specified and can be used to estimate the style groupings. Eq. (1) gives further insight into the nature of time-varying portfolio strategies, although the parameters in this equation are not identified without further restrictions.[1]

[1] A restriction sufficient for identification purposes is to assume that the portfolio strategy is constant over a number of months greater than the number of factors. This might seem unduly restrictive. However, for the purposes of characterizing the time-varying strategies of each style it suffices that we assume a quarterly holding period with two factors given by the return on cash and on equity investments. Other quarterly factors are captured in the α_{Jt} terms. Monthly data will then suffice to estimate Eq. (1). This approach can be contrasted with Sharpe's (1992) use of a rolling regression technology. The monthly updated portfolio shares should be interpreted as the average style-based portfolio shares for the previous 24 months.

In order to implement our style classification (SC) algorithm, we prespecify a number of styles.

A modification of the basic algorithm is a generalized least squares procedure (GSC), which allows time-varying and fund-specific residual return variance. By scaling observations by the inverse of the estimated standard deviation, we decrease the influence of extreme observations in the classification process. Amihud, Christiansen, and Mendelson (1992), for instance, find that this shrinkage improves forecasts of security returns. The details of the GSC procedure are provided in the Appendix.

It is important to note that the SC and GSC procedures make minimal demands on the available data. We can estimate Eq. (2) without needing to know factor loadings or style attributes represented by the vector β_{Jt} which may well change from period to period. We only need to know ex post returns on individual funds. There is a direct analogy between our estimation technique and cluster analysis procedures. The criterion being minimized in Eq. (2) via the SC algorithm is the same criterion applied in the k-means clustering approach (e.g., Hartigan, 1975). Cluster analysis usually attempts to minimize the squared differences within groups of k characteristics. In this context, the characteristics might include risk exposure and the features of the average stock in the fund portfolio. In our classification procedure, the k characteristics are month-by-month returns and the group means are the conditional expectations appearing in Eqs. (1) and (2). These characteristics are explicitly time-variant and capture not only risk but also dynamic portfolio strategies that are specific to particular fund styles.

Because we relax the requirement of constant portfolio weights through time, we would not expect to identify perfect analogues to the categories derived using a fixed linear time-series model, or even a very rich set of cross-sectional characteristics, including stock portfolio composition, observed at one point in time. Nonetheless, after estimating categories based on our classification algorithm, we report cross-tabulations with Morningstar mutual fund data fields, and also use the Sharpe (1992) procedure for estimating approximate fixed-positive-weight portfolio analogues using a standard set of wide-spanning asset class returns provided by Ibbotson Associates. These two procedures give some intuition about the resulting mutual fund clusters. Our comparisons yield evidence of different management strategies, which relate to known classifications such as 'growth' and 'value' management.

3.2. *How many styles?*

Because the procedure relies on prespecifying the number of styles, it is natural to ask what is the right number. To address this question, we use a likelihood ratio test suggested by Quandt (1960) for each successive decrease in the number of prespecified styles from nine. The test statistic for K styles (as opposed to $K + 1$) styles is

$$LR = Tm\left(\ln\frac{ssq_k}{Tm} - \ln\frac{ssq_{K+1}}{Tm}\right),$$

where T is the number of time periods, m the number of funds, and ssq_K and ssq_{K+1} are the appropriate heteroskedasticity-adjusted sum of squared errors. This statistic should be approximately distributed as χ^2 with $2T$ degrees of freedom.

Applying this measure to successive levels of fund aggregation, we find evidence for using at least eight separate categories. There is some ambiguity about the appropriate degrees of freedom, as well as the appropriateness of the χ^2 distribution in this case (see Quandt, 1960). Nonetheless, the observed test statistics are very large. For $k = 8$ through $k = 3$ styles, the test statistic values are 4,682.9, 4,092.1, 32,217.3, 6555.5, 7,106.2, and 10,197.7, respectively. In each case, the p-values are arbitrarily close to zero, indicating that an increase in the number of styles is useful in explaining returns. This result is similar to that reported for χ^2 tests for the number of factors, where typically too many factors are identified (e.g., Brown, 1989). An important caveat is that the χ^2 test is sensitive to departures from normality.[2] Using fewer than five groups, the distribution of the group returns suggests that the χ^2 test is well-specified. For these low numbers of groups, the algorithm clearly forces disparate funds together, increasing the model error. When the number of groups is increased beyond five, it is difficult to judge the relative magnitude of incremental improvement, although the sign of the test is positive for all values below nine, suggesting that more groups are needed.

3.3. Comparing procedures

A key question is how the GSC classification performs relative to standard industry classifications by investment objective. As Trzcinka (1995) points out,

[2] There are significant differences in skewness and kurtosis by style category (for a normal distribution, skewness is 0 and kurtosis is 3):

Group	Skewness	Kurtosis
1	0.0172	2.65
2	0.0540	3.45
3	0.1033	3.95
4	0.1481	3.97
5	−0.0670	5.37
6	0.2183	5.32
7	0.1697	5.52
8	0.1015	14.45
Entire sample	0.0782	4.11

The last four groups show significant departures from normality. Thus, the χ^2 distribution may be inappropriate for evaluating the unusualness of the test statistic. In other words, gains to increasing the number of styes above five groups may be overstated.

there are no generally accepted standards for comparing style classifications, which are put to a broad range of uses, from developing benchmarks for risk and return to establishing specifications used in investment management contracts. Despite this ambiguity, we borrow a natural measure from the asset pricing literature. Specifically, we compare our empirically determined styles with the classifications provided by Weisenberger over the period 1976 through 1992 and Morningstar over the period 1993 through 1995. The reason we use Weisenberger for the early period rather than Morningstar is that mutual fund styles change through time. The Weisenberger style codes were obtained at the end of each year in the sample period, and thus they have no ex post bias. Funds are classified using the GSC algorithm applied to data up to and including the Weisenberger publication date, with the number of styles chosen to match the number of industry objectives extant in the last month of the estimation window. Fund returns are then computed over the following year. Results are qualitatively similar using a one-month test period and using rolling month-by-month returns for 24, 36, 48, and 60 months to classify funds. However, the performance of the industry-based style categories is notably inferior to the other methods. This is not surprising, since the other methods use data subsequent to the publication date of the industry styles to classify funds. For this reason, we report results using a one-year test period.

We then cross-sectionally regress fund returns on a matrix of dummy variables that indicate whether each fund belongs to a particular style. If the style classification contains information about future differences in returns, we would expect these regressions to explain a significant amount of cross-sectional variation. This same procedure is performed for the industry classifications. A comparison of adjusted R^2s indicates which has the superior predictive ability. This procedure resembles classical time-series, cross-section tests of pricing models, except that the cross-sectional regressors are not loadings but a matrix of dummy variables.

As an alternative to the classification based on returns, we report a variety of other reasonable classification schemes. First, we classify funds based on latent variable factor loadings derived from a principal components analysis applied to the time-series matrix of fund returns in the estimation period. We apply the classification algorithm to the fund loadings on these factors. This is analogous to the principal component reduction in Elton and Gruber (1970), who use loadings as the inputs to a classification algorithm. We estimate the loadings following the procedures described in Connor and Korajczyk (1986) and Lehmann and Modest (1987).

Second, we prespecify factors and use the prespecified factor loadings to apply the SC algorithm. This approach has extensive precedent in the empirical literature. Chen, Roll, and Ross (1986) and Berry, Burmeister, and McElroy (1988), for example, prespecify macroeconomic risk factors to analyze stock portfolios and industry characteristics. In an application to mutual funds, Lehmann and Modest (1987), Grinblatt and Titman (1989), Elton, Gruber, Das, and Hvakla (1993), and Hendricks, Patel, and Zeckhauser (1993) all prespecify 'control' portfolios according to factors such as size and dividend yield. There are numerous other examples. For the purposes of this analysis, we use eight indexes: gold, the EAFE minus U.S. global equity index, the EAFA European equity index, the EAFA Pacific equity index, U.S. Treasury bills, commercial paper, long-term government bonds, long-term corporate bonds, high-yield bonds, the S&P 500, small stocks, and IPO's, all obtained from Ibbotson Associates. This approach has several advantages. First, the profile of each category has some intuitive interpretation – one group may be tilted towards bonds, while another is tilted towards stocks, for instance. Second, it suffers less from the difficulty of heteroskedasticity across funds that introduces systematic error into the endogenously determined principal components. Third, the coefficients (when properly scaled) have a natural interpretation as portfolios. The drawbacks are, of course, that the procedure does not allow for temporal variation in the portfolio weights. Finally, we cluster in the space of 'Sharpe coefficients' (see Sharpe, 1992) estimated on the same capital market indexes as above. These are estimated via a constrained optimization procedure under the assumptions that the weights remain fixed over the estimation period, that they are nonnegative, and that they sum to one. The weights can thus be interpreted as portfolio weights for passive, investable indexes.

As a benchmark to the performance of these various classification alternatives, we also report cross-sectional regression results for the factor loadings themselves. In other words, we use as independent variables in the cross-section regression four separate sets of explanatory variables. First, we use the coefficients estimated for each fund obtained by regressing the individual fund return series on the set of SC styles. Second, we use the first k principal components (where k corresponds to the number of extant industry objectives). Third, we use the capital market indexes described above. Fourth, we use the Sharpe coefficients. These four alternate procedures allow us to quantify how much is lost by reducing the continuous coefficients down to a simple classification scheme.

4. Empirical results

4.1. Summary of GSC categories

In Table 1, we report the cross-tabulation of the GSC categories with the Morningstar categories. Since Morningstar categories are identified only at one point in time, i.e., at the end of the sample period, we would not expect a perfect correspondence. The key feature of Table 1 is that the 'growth' category, which is the single largest designation for Morningstar, is spread widely across several different GSC categories, especially categories 1, 3, 4, and 5. While it is common

Table 1
Cross-tabulation of equity funds by Morningstar and GSC categories, summary of results using GSC algorithm, January 1976 to December 1994

GSC group	1	2	3	4	5	6	7	8	Total
Equity income	23	3	74	5	0	0	0	0	105
Europe	0	0	2	1	17	0	17	0	37
Foreign	0	0	2	2	137	1	136	0	278
Growth	196	296	54	117	1	77	0	0	741
Growth income	247	34	89	21	0	0	0	0	391
Pacific	0	0	0	0	42	0	28	0	70
Small-cap	0	16	7	132	0	115	1	0	271
Financial	1	0	5	7	2	0	0	0	15
Health	2	7	0	0	0	7	0	0	16
Metal	0	0	0	0	1	0	0	35	36
Nat.Resource	0	8	1	9	0	6	1	32	
Technology	0	7	0	3	1	19	0	0	30
Unaligned	2	5	4	24	1	0	0	0	36
Utilities	1	2	75	1	3	1	0	0	83
World	3	3	12	4	85	7	27	1	142
Total	482	373	332	318	299	227	215	37	2283

The table reports the cross-tabulation of mutual fund GSC categories with Morningstar style categories. The Morningstar categories are those attributed to the funds as of 1994 by the company itself, and thus do not take into account style shifts through the sample period. The unaligned group includes miscellaneous sector funds, such as REIT funds. The GSC procedure is a maximum likelihood method described in the text. It allows portfolio weights to vary on a quarterly basis, with eight factors pre-specified. A likelihood ratio test suggested by Quandt and Ramsey (1978) shows that the cross-section of mutual fund returns are driven by at least eight separate factors, for which loadings may vary.

to approximately control for risk in mutual fund studies by focusing only on growth funds (see, for example, Hendricks, Patel, and Zeckhauser, 1993; Brown and Goetzmann, 1995; and Ibbotson and Goetzmann, 1994), Table 1 indicates that many different portfolio strategies can fall under that broad rubric. Indeed, the GSC algorithm groups a significant percentage of growth funds with growth-and-income funds, suggesting that these labels do not provide particularly useful distinctions for investors. Also note that the small-cap category splits into two distinct groups. Apparently, the average capitalization of the stocks in the portfolio is not a sufficient statistic for performance. For the sector funds, the Morningstar classifications and GSC classifications generally agree. The health, metals, utilities, and unaligned (possibly real estate) categories are unambiguous. The technology sector and natural resource sector (which includes forest products as well as oil and gas) are split. It is clear from Table 1 that GSC group 8 is the precious metals fund category: it includes no funds other than metal sector funds. It also appears that group 1 is composed mostly of growth-and-income funds, group 2 is composed mostly of growth funds, and most utility sector funds fall into group 3, suggesting that this third group is an equity-income category.

Table 2 provides further insight into the characteristics of the GSC categories. For each category, we estimate the mean and standard deviation of portfolio weights, assuming a 24-month (nonoverlapping) return interval. Following Sharpe (1992), we constrain the coefficients to be positive, so that they can be interpreted as weights in short-sale constrained analogue portfolios. Groups 1 and 2 have a large average exposure to the S&P 500, groups 4 and 6 have a large exposure to the small-company stock index, and groups 5 and 7 have large exposures to non-U.S. indexes. Group 8 has a large exposure to gold.

Table 3 provides further evidence on the dynamics of manager strategies. We decrease the nonoverlapping estimation interval to six months in order to pick up variations in exposure to key indexes. In addition, we estimate the correlation of the style return to the previous period's index return. Thus, positive correlations indicate 'trend-chasing', while negative coefficients indicate a 'contrarian' approach. Groups 1 and 3 both have negative portfolio weight correlations to lagged S&P 500 return values. Groups 5, 6, and 7 all have positive portfolio weight correlations to lagged S&P 500 return values. Group 7, an international style, is most heavily weighted towards the EAFE index and is little invested in the U.S. market. Group 7 managers tend to buy the EAFE stocks when returns were low last period. Group 5, the other international style, has much greater weight on the S&P 500, and appears to be a 'trend chaser' with respect to the U.S. market.

Table 4 reports cross-tabulations of Morningstar 'On-Disc' categories with GSC and Morningstar groups. It reveals useful information about fund strategies. For instance, it indicates average P/E ratios, average price-to-book ratios, and average ex post five-year earnings growth. These measures have been found to explain differences in security returns. They also appear to explain differences in style classifications. While these data represent a snapshot of the funds as of the last date in our data base, we believe they provide an important validation for the style classification procedure.

There are two styles that are composed of Morningstar 'small-cap' funds. Group 4 managers invest in stocks with low price-to-book ratios and low price–earnings ratios. These are 'value' managers. Group 6 managers invest in companies with high price-to-book ratios and high ex post earnings growth. These are 'glamour' stock managers who purchase companies that have grown rapidly in the past. Group 4 managers buy stocks with low betas, and group 6 managers buy stocks with high betas. Group 4 managers buy financials, cyclicals, and services, and Group 6 managers buy health care stocks and high technology issues. In the terminology of Lakonishok, Shleifer, and Vishny (1994), these are 'glamour' managers. In view of the behavioral model proposed

Table 2
Mean and standard deviation of 24-month (nonoverlapping) Sharpe implied portfolio weights, December 1978 to December 1994

	IPO	Small	S&P	Junk	Lt corp	Lt gvt	Cpaper	T bills	Gold	Eafe	Europe	Pacific
Group 1												
Mean	0.00088	0.14190	0.75792	0.0072	0.0132	0.0172	0.0060	0.0273	0.0134	0.0000	0.0092	0.0062
Std. dev.	0.00121	0.04295	0.05583	0.0112	0.0156	0.0258	0.0130	0.0304	0.0115	0.0000	0.0145	0.0131
Group 2												
Mean	0.00192	0.29909	0.67357	0.0000	0.0000	0.0051	0.0000	0.0000	0.0127	0.0000	0.0066	0.0018
Std. dev.	0.00274	0.07125	0.09463	0.0000	0.0000	0.0154	0.0000	0.0000	0.0324	0.0000	0.0131	0.0055
Group 3												
Mean	0.00172	0.12419	0.48548	0.0480	0.0795	0.0282	0.0843	0.0977	0.0211	0.0176	0.0090	0.0039
Std. dev.	0.00220	0.06922	0.05546	0.0552	0.0975	0.0245	0.0977	0.0832	0.0209	0.0283	0.0100	0.0103
Group 4												
Mean	0.00388	0.57963	0.36383	0.0000	0.0083	0.0056	0.0000	0.0000	0.0190	0.0000	0.0210	0.0004
Std. dev.	0.00392	0.08008	0.06691	0.0000	0.0135	0.0168	0.0000	0.0000	0.0237	0.0000	0.0305	0.0007
Group 5												
Mean	0.00130	0.22508	0.18613	0.0068	0.0070	0.0178	0.0511	0.0149	0.0292	0.1509	0.2796	0.0307
Std. dev.	0.00143	0.09681	0.24990	0.0180	0.0211	0.0342	0.0871	0.0385	0.0427	0.1132	0.2172	0.0467
Group 6												
Mean	0.00462	0.61241	0.35686	0.0000	0.0000	0.0000	0.0000	0.0000	0.0177	0.0000	0.0105	0.0000
Std. dev.	0.00593	0.13444	0.17015	0.0000	0.0000	0.0000	0.0000	0.0000	0.0510	0.0000	0.0315	0.0000
Group 7												
Mean	0.00058	0.06468	0.03004	0.0140	0.0212	0.0000	0.0240	0.0149	0.0154	0.2758	0.2187	0.3209
Std. dev.	0.00130	0.06293	0.06069	0.0278	0.0387	0.0000	0.0605	0.0385	0.0243	0.3090	0.1875	0.3010
Group 8												
Mean	0.02292	0.10274	0.00332	0.0000	0.0000	0.0000	0.0000	0.0000	0.7557	0.0364	0.0466	0.0425
Std. dev.	0.02483	0.14125	0.00997	0.0000	0.0000	0.0000	0.0000	0.0000	0.1952	0.1091	0.0713	0.0880

The table reports summary statistics about the time-series of Sharpe coefficients, i.e., implied portfolio weights calculated using the procedure in Sharpe (1992) for each GSC style over the period 1976 through 1994. Coefficients are constrained to be constant over nonoverlapping 24-month periods.

Table 3
Mean standard deviation and trading correlations of 6-month (nonoverlapping) Sharpe implied portfolio weights, December 1978 to December 1994

	S&P	T bills	EAFE-US
Group 1			
Mean	0.88868	0.07536	0.03596
Std. dev.	0.06872	0.06268	0.04653
Corr.	− 0.36285	− 0.11794	0.08072
Group 2			
Mean	0.92767	0.02873	0.04361
Std. dev.	0.14101	0.07967	0.07964
Corr	0.00976	0.02262	− 0.04144
Group 3			
Mean	0.65325	0.27648	0.07027
Std. dev.	0.12637	0.12772	0.07421
Corr	− 0.33967	− 0.05222	0.07570
Group 4			
Mean	0.80581	0.09788	0.09632
Std. dev.	0.21447	0.16608	0.15136
Corr.	− 0.04370	− 0.08677	− 0.08074
Group 5			
Mean	0.53262	0.13748	0.32990
Std. dev.	0.29663	0.18198	0.19950
Corr.	0.33378	0.05611	− 0.04871
Group 6			
Mean	0.90455	0.02557	0.06988
Std. dev.	0.20093	0.09784	0.13963
Corr.	0.18407	0.04409	0.02801
Group 7			
Mean	0.10346	0.14319	0.75335
Std. dev.	0.12823	0.15949	0.18585
Corr.	0.25193	0.10338	− 0.26766
Group 8			
Mean	0.17685	0.41080	0.41235
Std. dev.	0.30928	0.41667	0.40264
Corr.	0.25022	− 0.00200	− 0.06077

This table reports the summary statistics for the time-series of Sharpe coefficients, i.e., implied portfolio weights calculated using the procedure in Sharpe (1992) for each GSC style over the period 1976 through 1994.

Coefficients are constrained to be constant over rolling 6-month periods. EAFE-US is the EAFE index of global equity returns not including the U.S. market. Correlation is between change in portfolio position and previous period index return. Change in portfolio position is measured as (− 1, 0, + 1) relative to the previous semiannual portfolio bought and held into the current period.

by these authors, it is not surprising to find that these glamour managers are also 'trend chasers' as is evident from the results of the previous table, engaging in almost twice the amount of trading of their 'value' counterparts in group 4.

Table 4
Cross-tabulation of Morningstar performance fields with Morningstar classifications and GSC categories

Panel 1: Average net asset value in by Morningstar and GSC categories

Morningstar categories	Generalized stylistic classification categories 1	2	3	4	5	6	7	8	Total
Equity income	233.15	15.87	582.96	24.54	NA	NA	NA	NA	463.54
Europe	NA	NA	10.55	3.00	230.75	NA	116.69	NA	160.28
Foreign	NA	NA	151.70	365.95	308.48	1.40	232.43	NA	269.15
Growth	240.36	345.51	174.74	616.20	9.80	356.39	NA	NA	348.60
Growth income	620.84	240.25	529.04	204.59	NA	NA	NA	NA	546.19
Pacific	NA	NA	NA	NA	195.03	NA	187.86	NA	192.08
Small company	NA	176.73	41.9	180.11	NA	192.34	41.80	NA	180.98
Finance	222.20	NA	151.58	208.13	3.10	NA	NA	NA	162.88
Health	457.30	341.56	NA	NA	NA	231.34	NA	NA	307.81
Metals	NA	NA	NA	NA	44.00	NA	NA	132.44	129.98
Nat. resource	232.43	NA	108.74	183.90	43.67	NA	19.52	20.20	100.35
Technology	NA	197.71	NA	314.47	176.40	183.10	NA	NA	199.42
Unaligned	113.10	27.24	185.00	52.89	81.60	NA	NA	NA	68.15
Utility	369.60	5.30	299.40	37.60	857.03	17.30	NA	NA	306.85
World	1,715.80	76.17	554.78	1,206.10	290.49	61.77	127.72	2.00	319.95
Total	445.51	314.96	396.76	344.68	275.51	242.49	197.17	125.88	332.23

Panel 2: Average value of fund characteristics by GSC categories (NAV weighted)

Morningstar categories	Generalized stylistic classification categories 1	2	3	4	5	6	7	8	Total
Alpha	0.82	0.52	1.75	2.25	2.4	0.51	3.41	11.88	1.32
Beta	0.92	1.00	0.77	0.94	0.78	1.11	0.58	0.30	0.89
Rsquared R^2	82.91	67.26	67.14	60.76	26.93	41.63	13.33	1.10	62.73

Turnover	36.47	92.92	65.26	122.49	48.64	103.23	47.24	29.60	76.46
P/E ratio	18.07	19.92	18.79	20.2	25.41	24.00	24.21	29.64	20.52
Price/book ratio	3.47	3.64	2.40	3.21	3.22	4.29	2.37	2.54	3.31
5-Year earnings growth	4.90	15.31	5.28	9.92	1.45	23.59	1.91	0.36	5.68
Return on assets	7.36	8.63	4.77	8.2	7.37	10.83	6.31	11.38	7.59
Debt to capital	30.19	27.30	34.00	29.52	28.90	24.81	26.43	19.76	28.37
Median market cap.	9,144.4	5,194.5	4,170.1	1,415.5	5,868.1	820.8	2,850.9	1,255.1	5,613.6
Energy (%)	10.72	5.27	NA	NA	NA	NA	NA	NA	8.51
Financials (%)	18.61	13.88	19.95	15.50	21.31	6.97	17.69	2.66	16.84
Industrial cyclicals (%)	17.53	NA	NA	NA	NA	NA	NA	NA	17.53
Consumer durables (%)	6.83	7.76	5.33	8.63	10.69	6.77	12.01	0.20	7.63
Consumer staples (%)	7.11	5.62	5.42	3.27	5.57	1.98	5.55	0.03	5.36
Services (%)	9.89	13.55	7.94	13.77	11.9	12.09	10.19	1.28	11.02
Retail (%)	4.81	6.07	3.53	6.01	4.23	7.29	5.96	0.09	5.11
Health (%)	8.53	11.36	6.34	6.95	4.06	15.33	4.39	0.02	8.09
Technology (%)	9.28	20.82	3.97	22.75	5.94	36.22	6.59	0.06	13.48
Cash (%)	6.38	12.08	9.86	7.75	13.02	9.00	5.71	5.35	8.92
Equity (%)	89.57	84.31	75.08	90.16	83.38	90.10	91.79	91.96	85.83
Bonds (%)	2.73	1.95	9.64	1.35	1.74	0.30	1.31	0.12	3.23
Preferred (%)	0.12	0.40	0.79	0.09	0.91	0.27	0.22	0.25	0.38
Other (%)	1.06	1.11	4.59	0.61	0.90	0.33	0.59	2.32	1.53
Foreign (%)	8.23	9.99	14.54	9.16	88.30	7.87	96.3	81.74	23.68
Sales charge	2.85	2.88	2.89	2.91	3.50	2.21	1.59	2.45	2.82
Front end load	2.48	2.55	2.09	2.75	2.89	1.78	1.09	2.16	2.37
Expense ratio	0.77	1.10	1.04	1.06	1.41	1.28	1.21	1.19	1.04

Panel 1 reports the cross-tabulation of fund net asset values in millions of dollars as of December 1994 for the GSC style categories (columns) and the Morningstar fund objective classifications (rows) as of that date. Abreviations for objective classifications are the same as in Table 1. NA indicates that summary data on net asset values was unavailable for that category.

Panel 2 reports the average value of a number of fund characteristics calculated by Morningstar Inc. as of December 1994 for each of the GSC style categories, as well as for the entire sample. 'Alpha', 'Beta', and 'R^2' are regression statistics estimated from a single-factor market model using the S&P 500 as the regressor over the history of the fund. (%) following a category indicates the average percentage of the fund portfolio invested in securities of that type.

Taken together, Tables 1, 2, and 3 suggest a categorization that is somewhat different from the typical industry groupings. The following summarizes our GSC styles:

Category 1: *'Growth and Income'* Comprised primarily of Morningstar 'growth' and 'growth-and-income' funds, Category 1 funds have the highest positive weights of any category on the S&P 500. They invest in relatively large companies.

Category 2: *'Growth'* Comprised primarily of Morningstar 'growth' funds, Category 2 funds have a major exposure to the S&P 500 and to a lesser extent to small stocks. Their exposure to debt asset classes is minor.

Category 3: *'Income'* Comprised primarily of Morningstar 'equity income', 'growth-and-income', and 'utility sector' categories, Category 3 funds have the highest cash balances and the highest exposure to debt asset classes.

Category 4: *'Value'* Comprised mostly of 'small-cap' funds, this category seeks stocks with low price-to-book and low price–earnings ratios. This is consistent with value management.

Category 5: *'Global Timing'* Comprised principally of non-U.S. equities, this category nonetheless pursues a dynamic strategy of increasing exposure to the U.S. market when it rises; the variability of this U.S. exposure suggests a timing strategy.

Category 6: *'Glamour'* Comprised primarily of Morningstar 'growth' and 'small company' funds, Category 6 funds have a major exposure to small-cap stocks. The equities in the typical portfolio have relatively high price-to-book and P/E ratios and high ex post five-year earnings growth. These are also domestic 'trend-chasers', displaying positive correlation to preceding S&P index returns.

Category 7: *'International'* Comprised of funds that are not strongly exposed to the U.S. market through time, Category 7 funds appear to vary their exposure to European and Pacific markets considerably.

Category 8: *'Metal Funds'* Comprised entirely of funds from the 'precious metals and commodities' Morningstar category.

4.2. Stability of styles

How consistent are the style classifications? Because of their statistical nature, all classification schemes run the risk of misclassification. Funds at the margin between two styles, for instance, may be difficult to confidently allocate to one style or the other. To address the problem of estimation error, we use a bootstrapping procedure to determine the frequency with which funds 'switch' classifications. Using a single 24-month window, we repeatedly apply the SC algorithm to the same data and count the number of changes in the pairwise associations between each fund in the sample. We find an average 'switching rate' of 11%.[3] We also bootstrap the switching rate under the null hypothesis of no cross-sectional structure. This null is constructed by forming 24-month samples via random draws without replacement from actual fund returns. The typical rate of change under this null is 27.3%.

The bootstrapped null distribution is helpful in examining the question of style stability through time. To do this, we apply the SC algorithm to rolling 24-month windows of mutual fund returns. As a measure of the style stability over successive 24-month windows, each year we count the number of changes in pairwise associations between each fund. The average annual percentage of fund associations that change each year is 17.6%. This is higher than the 11% we would expect due solely to statistical variation conditional on no classification change, but considerably less than we would expect under the null, i.e., if the classifications were spurious. In fact, we find that 12 of 16 of the sample years have percentage changes below the 5% quantile of the bootstrapped null distribution, allowing us to reject the null for that year. For further details of bootstrapping association frequencies see Abraham, Goetzmann, and Wachter (1994) and Goetzmann and Wachter (1995).

Our bootstrapping tests suggest that the GSC and SC algorithms manifest statistical variation. Like most statistical measures, they are noisy. Despite this noise, our tests reject the hypothesis that our management styles are spurious. On the other hand, the bootstrapping tests are silent on the usefulness of the resulting classification scheme. The true value of style classification rests in the extent to which it successfully explains out-of-sample fund performance. The next section addresses the usefulness of the SC and GSC classifications.

4.3. *Explanatory power of styles*

How well do estimated fund styles explain cross-sectional variation in out-of-sample fund returns? Table 5 reports the results of out-of-sample cross-sectional regressions for each of the categorization methods as well as for the industry

[3] Note that this measure overrepresents the percentage of funds that change classification each year. For example, take a sample of five funds, two of which remain in one category over two iterations, and two of which remain in the other category over two iterations. However, the fifth fund changes its association from the first group to the second group. This results is a change of association with each of the four other funds in the sample out of the possible ten pairwise associations – a 40% switching rate, generated by only 20% of the funds changing classification. When the number of funds in each of the two groups is equal, the misclassification percentage is half the swtching rate.

Table 5
Cross-sectional return variance explained by ex ante classification methods and factor loadings

Regressing returns on classifications: Adjusted R^2 regressing returns on factor loadings: Adjusted R^2

Test period	Return based classifications (GSC procedure)	Classifications based on Sharpe coefficients	Classifications based on principal components	Classifications based on style categories	Constrained pre-specified factors (Sharpe procedure)	Principal factor loadings	GSC centers	Unconstrained pre-specified factor loadings
1978	0.372	0.194	0.287	0.227	0.368	0.44	0.421	0.45
1979	0.359	0.050	0.331	0.234	0.459	0.526	0.533	0.522
1980	0.306	0.135	0.323	0.199	0.423	0.485	0.499	0.498
1981	0.381	0.151	0.388	0.205	0.484	0.518	0.521	0.521
1982	0.390	0.152	0.395	0.188	0.488	0.489	0.470	0.511
1983	0.299	0.057	0.241	0.124	0.308	0.287	0.295	0.304
1984	0.257	0.051	0.221	0.128	0.302	0.297	0.290	0.305
1985	0.177	0.033	0.108	0.038	0.262	0.222	0.224	0.239
1986	0.080	0.013	0.043	0.035	0.148	0.103	0.104	0.127
1987	0.177	0.033	0.131	0.093	0.179	0.174	0.174	0.172
1988	0.400	0.070	0.436	0.194	0.501	0.506	0.504	0.503

1989	0.374	0.066	0.362	0.192	0.421	0.425	0.428	0.426
1990	0.423	0.129	0.376	0.208	0.483	0.476	0.477	0.473
1991	0.351	0.085	0.368	0.193	0.402	0.406	0.411	0.427
1992	0.358	0.064	0.352	0.191	0.398	0.411	0.424	0.459
1993	0.174	0.032	0.166	0.122	0.167	0.177	0.184	0.188
1994	0.196	0.031	0.16	0.107	0.184	0.205	0.207	0.218
Mean	0.298	0.079	0.276	0.158	0.352	0.362	0.363	0.373
Median	0.351	0.064	0.323	0.191	0.398	0.411	0.421	0.427
Std. dev.	0.102	0.053	0.118	0.062	0.124	0.142	0.141	0.139

This table uses a 24-month estimation period prior to and including the Weisenberger publication date to estimate coefficients and form classifications. The number of classifications in each period is specified by the number of industry objective codes provided by Weisenberger (1978–93) and Morningstar (1994). The out-of-sample test period corresponds to the 12 months between Weisenberger publications. The cross-section of test period returns on funds are regressed against $(K - 1)$ dummy variables, where $\delta_{ki} = 1$ for fund I in category k and zero otherwise. The first column gives adjusted R^2 for the categories given by the GSC procedure described in the text, the second and third columns correspond to categories based on constrained Sharpe coefficients and principal factors procedures (using the SC procedure), while the fourth column uses the Weisenberger style categories (1978–93) and Morningstar categories (1994). These are compared to the adjusted R^2 obtained by using the Sharpe coefficients (Column 5), factor loadings (Column 6), loadings on the SC style centers (Column 7), and the unconstrained loadings on capital market returns used to estimate the Sharpe coefficients (Column 8). The data are total returns to U.S. equity mutual funds, excluding sector funds, but including international funds, over the period 1976 through 1994.

objective classifications. We omit sector funds from the analysis, since there is relatively little ambiguity about their classification. Instead of using the entire history of fund returns to form styles, we use the rolling period of 24 months for estimation purposes. As shown in the preceding section, this results in less 'stable' styles, but it relaxes the assumption that funds belong to the same style over the entire period and only uses ex ante industry information.

Columns 1, 2, and 3 in each panel show the adjusted R^2 that results from the application of the iterative relocation algorithm to different spaces: the space of returns, the space of 'Sharpe coefficients', and the space of principal component loadings. Column 4 reports the results based on the industry objective classification. Notice that, although adjusted R^2s differ for various estimation intervals, grouping in the space of returns and grouping in the space of factor loadings typically explain significant amounts of performance. Grouping funds according to the Sharpe coefficients performs about as well as using the industry codes. This may be due to the fact that, for any fund, a significant number of coefficients are zero, due to the nonnegativity constraint. For each estimation period, the GSC algorithm applied to returns marginally outperforms the algorithm applied to loadings on principal components. This may be due to the fact that the model of classification in Eq. (1) is well-specified when loadings change through time, but principal components rely upon stationary loadings. The GSC categories explain about one-third of cross-sectional variation of returns, ex ante. The Weisenberger categories explain, on average, 16% of the variation in fund returns, while classifying funds according to Sharpe coefficients explains only about 8%, on average.

The last four columns in each panel of Table 5 report the percentage of cross-sectional variation explained by the estimated factor loadings themselves. We would expect these to have greater explanatory power, since they are continuous rather than dummy variables. This is particularly important for outlying funds that have extreme exposures to some factor. While the GSC algorithm will either group this outlier by itself or lump it in with distant neighbors, the factor loadings themselves may capture the magnitude of its deviation in cross-section. In addition to using the Sharpe coefficients and principal component factor loadings as regressors, we also create indexes based on the SC centers in the space of returns, and estimate unconstrained loadings on passive indexes. This last column allows us to examine how much explanatory power the Sharpe procedure gives up in return for its positivity constraint. The Sharpe positivity constraint is useful, because it allows the coefficients to be interpreted as a vector of portfolio weights on investable indexes. Our style categories do not have this property. Consequently, our GSC procedure is not intended as a competing procedure to the Sharpe 'style analysis'. In this paper, we show that the two tools can be used together to identify common strategies among managers. The GSC procedure identifies aggregate behavior, and the Sharpe procedure helps interpret it as strategy.

The second panel of Table 5 shows that the loadings themselves all perform better than the classification indicators. Typically, they explain on average about 6 percentage points more cross-sectional variation out-of-sample. This suggests that, in absolute terms, factor loadings, however they are constructed, are a superior method of risk adjustment. On the other hand, the GSC procedure does not do badly on a relative scale. Grouping by manager style is not an alternative method for risk-adjusting manager returns. However, given that benchmarking by style is a common practice, our analysis indicates that there is not a great deal of information lost by using simple style classifications that are appropriately chosen.

It is interesting to note that the loadings on the GSC centers typically outperform the constrained loadings on prespecified financial indexes and do a little better than the loadings on principal factors. They do almost as well as the unconstrained loadings on the prespecified factors. It is tempting to conjecture that the 'glamour' vs. 'value' division in the styles is responsible for the success of the simple multinomial statistic, since this division may capture one of the fundamental factors found to be superior in out-of-sample tests on U.S. equities (see Fama and French, 1992; Lakonishok, Shleifer, and Vishny, 1994).

4.4. Interpretation

It is not surprising that the categories based on returns outperform the standard industry classifications. Categories like 'growth' and 'growth and income' represent an invitation to fund management gamesmanship. Once a fund is classified into a particular category, there is little incentive to pursue an investment strategy that will ensure that future fund performance will be close to the category average in the future. Sirri and Tufano (1992) and Goetzmann and Peles (1996) report evidence that mutual fund investors flock to superior performers in each fund category. Given this information, fund managers are not rewarded by maintaining strategies consistent with their industry classification. As a result, we find that the GSC categories formed via returns and loadings 'agree to disagree' with the standard industry classifications. It is evident that the industry classifications have relatively little power to explain differential fund performance.

It is also not surprising that classification based on returns typically equals or beats classification based on principal factor loadings for longer holding periods. The principal component loadings represent a linear projection of the monthly returns on a reduced space. Both the reduction of dimensionality and the linearity of the projection represent constraints that 'coarsen' the information about fund returns. The advantage of the classification based on scaled principal components is the natural interpretation of the groups in terms of systematic risk classes. The disadvantage is that funds are subject to misclassification due to nonlinear strategies.

It is likewise not surprising to find that loadings on principal components outperform loadings on prespecified asset series. The principal components were selected so as to maximally spread returns in the preceding period. While Chen, Roll, and Ross (1986) show that prespecified factors likewise spread returns, collinearities among factors appear to increase the standard error of the factor loadings, and thus make fund classification via the SC procedure more difficult. The major advantage to estimating positive-constrained coefficients on prespecified factors is that it provides some insight into the composition and behavior of the categories.

Our results provide mixed signals regarding the approach to characterizing funds according to their profile of loadings on prespecified indexes. Even when the loadings are not constrained to be positive, we find no evident advantage in term of explanatory value. When loadings are used to identify styles, they perform as poorly as the standard industry classifications. Thus, their incremental advantage obtains in circumstances when the loadings themselves, rather than a derived style classification, can be used. Their disadvantage is that loadings on correlated indexes will be estimated with inaccuracy, due to collinearities.

5. Conclusion

We show that a simple procedure based on the switching regression technology applied to monthly returns dominates other management style classifications based on standard investment objectives and also does better than classification based on observed factor loadings. Given the potential 'category gaming' by fund managers, it is useful to have a classification method that uses publicly available and independently verifiable time-series information about returns.

Why are we interested in manager styles at all? Our classification algorithm identifies a few major types of fund strategies. While these may not exhaust the range of different fund managers, they do provide an overview of the strategies that differentiate managers. Our results validate the use of traditional, self-reported categories such as equity income, growth and income, and growth. However, funds apparently do not always correctly categorize themselves. We find two somewhat surprising divergences from standard industrial categories. First, we find evidence of 'value' vs. 'glamour' managers, rather than a monolithic small-cap group. Second, we find evidence that style involves dynamic strategies, rather than simply fixed portfolio weights.

The focus of this paper is the development of a classification algorithm that is consistent with commonly used asset pricing models. Such classification is unlikely to replace the use of continuous, multivariate models for risk adjustment, nor should it. The absolute magnitude of systematic risk exposures will always be important to portfolio decisions. Our analysis suggests, however, that there are a few intentional, recognizable strategies within the population of investment managers. As the finance profession continues to investigate the behavioral basis for investment decisions, it will be useful to further identify and study these common patterns. The GSC algorithm is a potentially useful tool for doing so. Its advantage over heuristic classification is that researchers can use it to decompose 'styles' into more familiar measures such as time-varying factor loadings and risk premiums.

Appendix: GSC methodology

The GSC can be thought of as an extension of a standard iterative relocation algorithm such as k-means. It is motivated by the insight that if $r_{it} = \mu_{it} + e_{it}$ and $\mathrm{var}(e_{it}) = \sigma_i \sigma_t z_{it}$, where z_{it} is i.i.d. normal for both I and t with mean 0 and variance 1, then $\mathrm{var}(e_{i.}) = \sigma_i^2 \mathrm{E}(\sigma_t)^2$ and $\mathrm{var}(e_{.t}) = \sigma_t^2 \mathrm{E}(\sigma_I)^2$, where we interpret σ_τ and σ_I as independent and identically distributed drawings from a population of time-series and cross-sectional standard deviations respectively independent of z_{it}. This is not inconsistent with a variety of GARCH or other processes for returns. As a result, $\mathrm{var}(e_{it})$ is proportional to $\mathrm{var}(e_{i.}) * \mathrm{var}(e_{.t})$. Therefore, it is a simple matter to infer the variance of each time and fund residual as proportional to the marginal time and fund variances in excess of the estimate of μ_{It}. Since the efficient (GLS) estimate of μ_{It} depends on σ_i, we need a second pass (GLS) for efficient estimation of both μ_{It} and $\mathrm{var}(e_{it})$.

Computationally, we proceed as follows. For a given definition of the clusters, we calculate

$$\hat{\mu}_{It} = \frac{\sum_{i \in I} R_{it}}{\mathrm{count}(i \in I)}.$$

Once this is done, we compute

$$\hat{e}_{it} = R_{it} - \hat{\mu}_{It}.$$

For all I, we then calculate $\mathrm{var}(\hat{e}_{i.})$, and for all t we calculate $\mathrm{var}(\hat{e}_{.t})$. Numerically, these numbers tend to be small, so we normalize them by the average marginal variances.

We now do a GLS correction for the mean, computing.

$$\hat{\mu}^*_{It} = \sum_{i \in I} \frac{R_{it}}{\mathrm{var}(\hat{e}_{i.})} \Big/ \sum_{i \in I} \frac{1}{\mathrm{var}(\hat{e}_{i.})}.$$

We use this updated GLS estimate of the mean to update variance measures. We also use this formula to update centroid means whenever funds are switched from one cluster to the next, although for computational simplicity we do not update variance measures at each switch. Denote the clusters formed at the jth switch as $I(j)$. Then the criterion function at the jth switch is proportional to

$$SSQ_j = \sum_{t=1}^{T} \sum_{I \in I_j} \sum_{i \in I} \frac{(R_{it} - \hat{\mu}^*_{It})^2}{\text{var}(\hat{e}^*_{i.})\text{var}(\hat{e}^*_{.t})}$$

using the result that $\text{var}(e_{it})$ is proportional to $\text{var}(e_{i.})^*\text{var}(e_{it})$.

References

Abraham, Jesse, William N. Goetzmann, and Susan Wachter, 1994, Homogeneous groupings of metropolitan housing markets, Journal of Housing Economics 3, 186–206.

Amihud, Yakov, Bent Christiansen, and Haim Mendelson, 1992, Further evidence on the risk return relationship, Working paper (Stern School of Business, New York University, New York, NY).

Berry, Michael, Edwin Burmeister, and Marjorie McElroy, 1988, Sorting out risks using known APT factors, Financial Analysts Journal 44, 29–42.

Brown, Stephen J., 1989, The number of factors in security returns, Journal of Finance 44, 1247–1262.

Brown, Stephen J. and William Goetzmann, 1995, Performance persistence, Journal of Finance 50, 679–698.

Brown, Keith C., W.V. Harlow, and Laura T. Starks, 1993, Of tournaments and temptations: Analysis of managerial incentives in the mutual fund industry, Working paper (University of Texas, Austin, TX).

Carleton, Willard and Victor McGee, 1970, Piecewise regression, Journal of the American Statistical Association, May, 1109–1124.

Chen, Nai-fu, Richard Roll, and Stephen Ross, 1986, Economic forces and the stock market, Journal of Business 59, 383–404.

Christopherson, Jon A., 1995, Equity style classifications, Journal of Portfolio Management 21, 32–43.

Connor, Gregory and Robert Korajczyk, 1986, Performance measurement with the arbitrage pricing theory: A new framework for analysis, Journal of Financial Economics 15, 373–394.

Connor, Gregory and Robert Korajczyk, 1991, The attributes, behavior and performance of U.S. mutual funds, Review of Quantitative Finance and Accounting 1, 5–26.

Donnelley, Barbara, 1992, What's in a name? Some mutual funds make it difficult to judge, Wall Street Journal, May 5, Section C, p. 1, col. 3.

Dybvig, Philip H. and Stephen A. Ross, 1985, The analytics of performance measurement using a security market line, Journal of Finance 40, 401–416.

Elton, Edwin, Martin Gruber, S. Das, and M. Hlavka, 1993, Efficiency with costly information: A reinterpretation of evidence for managed portfolios, Review of Financial Studies 6, 1–22.

Elton, Edwin and Martin Gruber, 1970, Improved forecasting through the design of homogeneous groupings, Journal of Business 44, 432–450.

Fama, Eugene and Kenneth French, 1992, The cross-section of expected stock returns, Journal of Finance 47, 427–466.

Ferson, Wayne and Rudi Schadt, 1996, Measuring fund strategy and performance in changing economic conditions, Journal of Finance 51, 425–461.

Goetzmann, William N. and Nadav Peles, 1996, Cognitive dissonance and mutual fund investors, Journal of Financial Research 20.

Goetzmann, William and Susan Wachter, 1995, Clustering methods for real estate portfolios, Journal of Real Estate Economics 23, 271–310.

Grinblatt, Mark and Sheridan Titman, 1988, The evaluation of mutual fund performance: An analysis of monthly returns, Working paper (John E. Anderson Graduate School of Management, University of California, Los Angeles, CA).

Grinblatt, Mark and Sheridan Titman, 1989, Mutual fund performance: An analysis of quarterly portfolio holdings, Journal of Business 62, 393–416.

Grinblatt, Mark and Sheridan Titman, 1993, Performance measurement without benchmarks: An examination of mutual fund returns, Journal of Business 66: 1, 47–68.

Grinblatt, Mark, Sheridan Titman, and Russ Wermers, 1993, Momentum strategies, portfolio performance and herding: A study of mutual fund behavior, American Economic Review 85, 1088–1105.

Hartigan, John A., 1975, Clustering algorithms (Wiley, New York, NY).

Hendricks, D., J. Patel, and R. Zeckhauser, 1993, Hot hands in mutual funds: The persistence of performance: 1974–88, Journal of Finance 48, 93–130.

Ibbotson, Roger and William Goetzmann, 1994, Do winners repeat? Patterns in mutual fund performance, Journal of Portfolio Management 20, Winter, 9–17.

Jensen, Michael, 1968, The performance of mutual funds in the period 1945–1964, Journal of Finance 23, 389–416.

Kim, Moon, Ravi Shukla, and Michael Tomas, 1995, Wolf in sheep's clothing: Do mutual fund objectives tell the truth?, Working paper (Syracuse University, Syracuse, NY).

Lakonishok, Josef, Andrei Shleifer, and Robert Vishny, 1994, Contrarian investment, extrapolation and investment risk, Journal of Finance 49, 1541–1578.

Lakonishok, Josef, Andrei Shleifer, Richard Thaler, and Robert Vishny, 1991, Window dressing by pension fund managers, American Economic Review 81, 227–231.

Lehmann, Bruce N. and David Modest, 1987, Mutual fund performance evaluation: A comparison of benchmarks and a benchmark of comparisons, Journal of Finance 42, 233–265.

Quandt, Richard, 1959, The estimation of the parameters in a linear regression system obeying two regimes, Journal of the American Statistical Association 53, 233–265.

Quandt, Richard, 1960, Tests of the hypothesis that a linear regression system obeys two regimes, Journal of the American Statistical Association 55, 324–330.

Quandt, Richard and James Ramsey, 1978, Estimating mixture of normal distributions and switching regressions, Journal of the American Statistical Association 73, 730–752.

Sharpe, William F., 1992, Asset allocation: Management style and performance measurement, Journal of Portfolio Management, Winter, 7–19.

Sirri, Erik and Peter Tufano, 1992, The demand for mutual fund services by individual investors, Working paper (Harvard Business School, Cambridge, MA).

Tierney, David E. and Kenneth Winston, 1991, Using generic benchmarks to present manager styles, Journal of Portfolio Management 17, Summer, 33–36.

Trzcinka, Charles A., 1995, Equity style classifications: A comment, Journal of Portfolio Management 21, Spring, 44–46.

Witkowski, Erik, 1994, Mutual fund misclassification: Evidence based on style analysis, B.A. thesis (Harvard University, Cambridge, MA).