# Objective Misclassification and Mutual Fund Performance

Dennis Bams [1]

Roger Otten [2]

Ehsan Ramezanifar [3]

Department of finance, School of Business and Economics, Maastricht University, The Netherlands

August 2015

**Abstract**

The actual investment objectives of mutual funds may be different from the stated objectives which are written in their prospectus. In this paper, we examine what percentage of U.S. mutual funds deviate from their stated investment objectives over time, highlight the driving forces of its behavior and investigate its impact on the relative fund performance. Using a Return-Based Style Analysis (RBSA), we find that 25 percent of the U.S. mutual funds differ from their stated investment objective and over 18 percent of the funds are severely misclassified. We report that on average misclassified funds are typically belong to high trading cost funds, high-expense funds, week prior performance and small-sized funds. We also find that funds that increase objective misclassification, perform worse than funds that adhere to their benchmark. The added value of the present study lies in gaining insight into the misclassified funds characteristics and the incentive of their behavior plus statistical tests for the RBSA by using a rich U.S. mutual funds database that was recently released by CRSP.

**Keywords:** Mutual funds, Style Analysis, Fund Performance, Misclassification, Nonlinear Optimization

# Introduction

Prior studies show that some mutual funds deviate from their stated investment style. If this deviation is detrimental to fund performance, investors are better off avoiding funds that are prone to changing their investment style over time. We document that on both an unadjusted and risk-adjusted basis, misclassified funds perform worse than well-classified funds. We next investigate the misclassified fund characteristics to answer a question what kind of mutual funds are more likely to deviate from their benchmarks. In sum, this paper first sheds more light on statistical features of Return Based Style Analysis by using the asymptotic distribution of the portfolio weights with wide U.S. equity mutual funds data from 2003 to 2014. Then we investigate the performance consequence of the investment objective deviation to assess the potential costs of misclassification that investors are exposed to. Our study also reports the misclassified fund characteristics. Although mutual fund style analysis has been the subject of extensive prior research, no attention has been paid to the potential reasons behind objective misclassification behavior and we attempt to fill this gap.

Mutual funds are required by regulation to express their investment style and policy in their prospectuses and they have to adhere to their stated objective. On the other hand, stated investment objectives are used by investors to choose a fund which fits with their risk and return profile. According to a survey that has been carried out by the Investment Company Institute, the national association of investment companies, 34% of fund investors read a fund's prospectus before purchasing a mutual fund[1]. This study also reports that knowing the fund investment objective before a purchasing mutual fund is important at least for 40% of fund investors. Although reading the fund's prospectus seems an obvious starting point, one of the pressing questions for investors is whether a fund's prospectus represents the actual investment

[1] ''Understanding Investor Preferences for Mutual Fund Information'' August 2006, the survey is based on the interviews with 737 randomly selected fund owners during five years.

style. If the stated investment objective is not the real objective that the funds pursue, it misinforms investors and leads them to the wrong investment decisions. Many studies such as Bartolomeo & Wikowski (1997), Brown & Goetzmann (1997), Kim, Shukla & Tomas (2000), Kim, White & Stone (2005), Cremers and Petajisto (2007), Andrew Mason et al. (2012) present evidence of a serious misclassification in investment objectives compared to actual styles. Therefore, getting more insight about the mutual fund investment style has been relevant to investors and regulators.

There are two popular approaches that are used to examine the investment objective misclassification. The first approach is the Holding Based Style Analysis (HBSA), and the second one is Return Based Style Analysis (RBSA). RBSA has some advantages in comparison to HBSA. For example, HBSA requires knowledge of mutual fund portfolio composition and hence it cannot be applied for many funds because the information is not available, at least not at the short-term frequency. Using HBSA can also lead us into the window-dressing problem, a practice in which fund managers distort their portfolios to mislead investors about the reality. Relevant studies include Lakonishok et al. (1991), Sias and Starks (1997), Meier and Schaumburg (2004) and Vikas Agarwal et al. (2014) that have provided evidence on window-dressing behavior in mutual fund industry. Conversely, RBSA relies only on mutual fund return data, which makes it an attractive tool. De Roon et al. (2004) and Rekenthaler et al. (2002) have carried-out an in-depth comparison of HBSA and RBSA and show that RBSA is more valid for predicting future returns than the study of historic composition of portfolios. Hence, in this study, we employ the method of Return Based Style Analysis.

Typically, in the RBSA method, the fund return is compared with the return on a number of selected passive style indices. The indices represent distinct investment styles within particular asset classes (e.g. value, growth, and small caps). Style analysis is the construction of a portfolio of indices that best mimics the historical performance of a mutual fund. The style of the fund is represented by the loadings on the indices.

Sharpe (1988, 1992) has proposed an econometric technique to conduct RBSA. This technique involves a non-linear constrained regression that uses several asset classes to replicate the historical return pattern of a portfolio. The regression analysis is proposed to arrive at point estimates for the portfolio weights. The ultimate idea is to check whether the estimated portfolio weights correspond to the targeted investment style of the mutual fund. The constraints are imposed to enhance an intuitive interpretation of the coefficients. First, to interpret the coefficients as weights within a portfolio the factor loadings are required to add up to one, henceforth called Portfolio Constraint. Second, coefficients should be positive to reflect the short-selling constraint. As Almazan, Brown, Carlson, and Chapman (2004) report, about 70% of mutual funds do not have the right to pursue any short selling activities and only 2% actually do sell short. Hence, most fund managers are subject to short-selling constraint henceforth called Positivity Constraint.

On the other hand, Sharpe style regression, which provides point estimates for the portfolio weights, only provides a limited picture of the information that is available in the historical fund returns. Therefore, determining a confidence interval for portfolio weights is more informative. Lobosco and DiBartolomeo (1997) propose approximate confidence intervals for the coefficients. This information is accurate in determining the preciseness of the style weights. These confidence intervals still provide limited information as they do not allow for tests on multiple coefficients. Kim et al. (2005) apply the methods of Andrews (1997a, 1999) to obtain an asymptotic distribution confidence interval and in addition they employ the result of Geweke (1986) to get a Bayesian confidence interval for the parameter estimation of a Sharpe style regression.

We propose an alternative approach to arrive at the asymptotic distribution of the parameter estimates. We include the entire asymptotic distribution of the style weights. These results are obtained by applying a combination of the Kuhn-Tucker optimization algorithm and standard bootstrapping. This allows us to infer confidence intervals for the style coefficients, and to carry out statistical tests on the parameters. In fact, we are considering the impact of parameter uncertainty on style analysis. Overall, we find that funds frequently have significantly different factor loadings than their benchmark. Moreover, we separate all funds into two groups henceforth called misclassified funds which deviate from their benchmarks and well-classified funds which adhere to their stated investment objectives. The result shows that over 25% of the mutual funds differ from their benchmarks and also more than 18% of them are severely misclassified.

After classifying mutual funds based on their deviation from the benchmarks, we next highlight the importance of objective misclassification on funds' performance. One possible reason to deviate from a stated investment style is objective gaming which means misclassified fund managers are deviating from their stated objectives to earn a higher relative performance. Although, fund managers change their investment objective style to take advantage from their market timing abilities, they may also expose investors to a higher risk-level than they aware of. Huang, Sialm & Zhang (2011) find that mutual funds that increase risk perform worse than funds that keep stable risk levels over time. They argue that risk shifting between mutual funds either is an indication of inferior ability or is motivated by agency problems.

Sirri and Tufano (1998), and Huang, Wei, and Yan (2007) find that there is a convex relation between funds' flows and their performance, that is mutual fund investors flock to superior performance in each fund category. Brown, Goetzmann, and Park (2001), Nanda, Wang, & Zheng (2004), Li & Tiwari (2006), Cremers & Petajisto (2009), Chapman, Evans, & Xu (2010), and Hu, Kale, Pagani, and Subramanian (2010) argue that the convex flow-performance relation can motivate mutual funds to strategically shift risk levels to get extra fund flows. Hence, because the real amount of funds' risks is unobservable as Kim et al. (2000) discuss, some funds are tempted to deviate from their stated objectives and take more risk to obtain a higher return. Therefore, active fund managers have are motivated to deviate from their stated investment objective to reach better performance and attract more fund flows. As long as this deviation leads funds to perform better than other funds, then it has a positive effect on investors. But if it exposes more risk to investors without sufficient return, it will be harmful to investors and should be prevented. To assess the potential costs of misclassification that investors are exposed to, obviously measuring and analysing the performance consequences of misclassification are of paramount importance. Overall, based on various fund performance measures, we find that misclassified mutual funds underperform the well-classified funds. Our empirical result suggests a negative relation between objective misclassification and fund performance.

As Carhart (1997), Cremers & Pareek (2014) and Huang, Sialm & Zhang (2011) discussed, mutual fund performance is negatively affected by the amount of trading that would be related to additional trading cost. Chalmers, Edelen and Kadlec (1999) describe that the turnover ratio captures a substantial fraction of trading cost. On the other hand, Odean (1998 and 1999) argue that one explanation for unnecessary turnover is overconfidence bias and alternatively, Dow and Gorty (1997) argue that some fund manager who are not overconfident, trade extremely to send positive signal to their investors and superiors to show that they are not passive in trading. Which means by increasing in delegated assets under management, the agency problem between asset owners and asset managers will be arise. This evidence show that excessive trading is hurting returns and leads funds to have poor performance, either caused by human biases or by the design of the delegated investment management industry. Hence, to study the impact of trading cost on fund performance we compare the turnover ratio of misclassified and well-classified mutual funds. We find that the turnover ratio of misclassified funds in growth and growth/income investment styles are significantly higher than in well-classified funds.

Therefore, poor performance of misclassified mutual funds are related to their additional trading cost.

Finally, style drift may be viewed as an example of the phenomenon described by Huang, Wei, and Yan (2007), who report that smaller funds have a more convex flow-performance relation, hence they might tend to deviate from their benchmarks more than others. We measure fund size based on their TNA and compare fund size of misclassified and well-classified funds. We find that misclassified funds, specifically in case of growth and growth/income investment styles have higher TNA than well-classified mutual funds.

Furthermore, Gil-Bazo and Ruiz-Verdu (2009) find that the high-expense funds perform worse than low-expense funds and suggest that these funds might target naive investors and are more tending to agency problem. Thus, high-expense funds may have bigger incentives to deviate from their stated investment objective to reach better performance and attract new investors. We find that higher-expense funds are more likely to deviate from their benchmarks, specifically in case of growth and growth/income investment styles. The results are also in line with Sensoy (2009) who find that mismatched self-designated benchmarks are more common among high-fee funds.

Cooper (2005) find that some mutual funds change their names to take advantage of current hot investment styles. Hence, based on the changes in superiority of indices over time, we divide our sample period into two equal sub-periods. We find that some fund managers intentionally deviate from their benchmarks because they invest in the “Hot” style and away from the current “Cold” style regardless of their stated investment objective.

Overall, we argue that objective misclassification is unlikely to be a signal of their market timing abilities. Instead, we argue that the underlying explanation is to be found in unskilled, overconfident or agency-prone funds managers. We also conclude that on average misclassified funds are typically belong to small-sized funds, weak prior performance, retail funds, high expenses and high trading cost than well-classified funds. Therefore, Misclassified funds are typically distressed funds that appear to deviate from their benchmarks to reverse current performance and investor redemption trends. And we argue that the deviation from stated investment style is zero game for fund managers.

To exemplify the application with respect to the fund misclassification phenomenon, we consider a wide sample of U.S. mutual funds that was recently released by CRSP. Following many relevant studies such as Jonathan Berk and Binsbergen (2014), Andrew Mason et al. (2012), Vikram Nada et al. (2009), Marcin Kacperczyk et al. (2008), we consider several selection criteria to arrive at the mutual fund data, relevant for our empirical study.

The remainder of the paper is organized as follows. In section 2 we introduce the methodology related to return-based style analysis and discuss the econometric technique to arrive at the asymptotic distribution. In section 3 we describe the data that are used in the empirical application. Section 4 contains empirical results, while in section 5 we address the robustness of these results. Finally, section 6 concludes the paper.

## 2. Methodology

The methodology used in our analysis has three features that are worth mentioning. First, we assess the preciseness of estimated style coefficients. Second, we test whether coefficients are significantly different from zero and third, we determine whether style coefficients are significantly different from each other. This extra information has important practical implications for the fund misclassification phenomenon and objective gaming. Incorporating the asymptotic distribution of style weights allows for statistical tests to track down misclassified funds and subsequently analyse their behavior. We apply this method to the fund misclassification phenomenon, considering a sample of U.S. mutual funds data. In section 2.1, we discuss the difference between two main approaches of fund style analysis and then in section 2.2 we concentrate on Return-Based Style Analysis to report statistical tests on Sharpe's model.

### 2.1. Return Based Style Analysis versus Holding Based Style Analysis

There are two widely accepted methods to perform mutual funds style analysis. The first is Return Based Style Analysis (RBSA) and the second is Holding Based Style Analysis (HBSA). HBSA in comparison with RBSA requires more knowledge about the fund holdings and the weight of different securities that it contains. In addition, to use HBSA, fund holdings information should be updated monthly, in order to get the latest data for the mutual fund, but in the United States, for example, the Security Exchange Commission (SEC) has forced fund managers to disclose their portfolio composition only twice a year[2]. Hence, the problem of data availability at least at a short-term frequency such as on a daily or monthly basis is the main drawback of the HBSA method. Therefore, if updated mutual fund holdings are not available, HBSA will lead to poor information. Furthermore, if complete information about the fund holdings is available, the methodology requires that the style characteristics of each security be identified (see for example Véronique Le Sourd (2007) and de Roon et al. (2004)). In addition, using HBSA can lead us into the window-dressing and a portfolio-pumping trap, a practice in which fund managers distort their portfolios to mislead investors about the reality. Relevant studies that have provided evidence on window-dressing behavior in mutual funds include Lakonishok et al. (1991), Sias and Starks (1997), Meier and Schaumburg (2004) and Agarwal et al. (2014).

Conversely, the RBSA requires minimal data (only returns) and low level of sophistication. Hence, RBSA has gained widespread attention among plan sponsors, investment consultants, and private investors. In this paper, we employ the RBSA method.

[2] At the end of its first and third fiscal quarters.

## 2.2. Stylized facts of Sharpe's model for return-based style analysis

The theory of RBSA asserts that a manager's investment style, both past and present, can be determined by comparing the manager's returns to the returns of a number of selected passive indices. Sharpe (1998) originally defined the RBSA approach. According to him, the principal goal is to find the best mimicking strategy that is in accordance with the investment style of the mutual fund. Sharpe (1998) proposes a factor model for RBSA that explains the returns for a given fund with the following model:

$$R_t = \alpha + \sum_{k=1}^{N} \beta_k I_{kt} + u_t \qquad t = 1, \ldots, T \tag{1}$$

where $R_t$ denotes the mutual fund return at time t, $N$ is the number of asset class factors, $\beta_k$ is a factor loading that expresses the sensitivity of the fund return to the factor-mimicking portfolio return of index $k$, $I_{kt}$ denotes the return of index k at time t and $u_t$ reflects idiosyncratic noise, orthogonal to the style indices, i.e. $E(I_{kt}u_t) = 0$.

Factor loadings have two main constraints. Firstly, they are restricted to add-up to one, in order to give them the interpretation of portfolio weights:

$$\sum_{k=1}^{N} \beta_k = 1 \tag{2}$$

Secondly, to meet the short-selling constraint that fund managers are mostly subject to, the following inequality constraints are imposed on the factor loadings:

$$\beta_k \geq 0 \qquad k = 1, \ldots, N \tag{3}$$

De Roon et al. (2004) have done an in-depth investigation about the effect of the portfolio and short-selling constraints in style analysis. They argue that although there is no straightforward analytical expression to define the benefit of imposing constraints, if both constraints are the case in reality, these results in more efficient parameter estimates. The latter is referred to as the strong form of RBSA.

According to equation (1), in this context $\sum_{k=1}^{N} \beta_k I_{kt}$ has the interpretation of the return on a passive portfolio with the same style as the fund. In the next section, we provide detailed information on the estimation algorithm for the factor loadings. Given parameter estimates for the factor loadings, the model in equation (1) subject to the constraints in (2) and (3) may have two applications: asset allocation and performance benchmarking.

Since the factor loadings have the interpretation of portfolio weights, RBSA is a useful tool to determine the asset allocation of the particular mutual fund. Besides examining the prospectus, talking to a fund's management and investment consultants, RBSA helps the investor to determine a fund's investment style.

If we interpret the estimated style weights as exposures to passive indices, return based style analysis is also applicable as a performance measurement tool. The return obtained by a fund in each month can be compared with the return on a mix of asset classes with the same estimated style. In equation (1), the systematic difference between the fund return and the estimated style index is represented by the intercept, $\alpha$. Because the input for return-based style analysis are indices which are; (1) a viable alternative (2) not easily beaten (3) identifiable and (4) easily replicated, the major criteria for measuring performance are met.

A crucial ingredient that may heavily influence the outcome of RBSA is the choice of the appropriate benchmarks. While Sharpe (1992) uses a detailed 12-asset class factor model, simpler models often yield more sensible results, for instance, in Lobosco and DiBartolomeo (1997).

A few prerequisites should be met before any reliable results are to be obtained. First, the benchmarks should be mutually exclusive which means they may not include any securities that already form part of any other basic asset classes considered in the model. Secondly, indices should be exhaustive benchmarks, meaning as many securities as possible should be included in the chosen asset classes. Thirdly, the correlation between returns on the basic asset groups considered in the proposed model should be low.

De Roon et al. (2004) argued that, the absence of linear combinations between the benchmarks for the explanatory asset classes included in the model assures that all of the three conditions are met. For instance, a mid-cap index is likely replicated by a weighted combination of a large cap index and a small cap index, and should therefore not be included. A way to control for this possible problem is to look at cross correlations and standard deviations. If correlations between specific benchmarks are too high, we could consider dropping some of them to diminish multicollinearity problems. The resulting model should be able to span the whole portfolio asset mix. Based on Vincent (2005), the analysis of multicollinearity among benchmarks can be done as follows:

(1) Calculation of Pearson’s correlation coefficients for the benchmarks considered.

(2) Calculation of Variance Inflation Factors (VIF) defines the degree to which each of the five benchmarks contributes to the multicollinearity of the model.

$$VIF = \frac{1}{1 - R_i^2} \qquad i = 1, \ldots, N \tag{4}$$

Where $R_i^2$ is the coefficient of regression when we regress index *i* on the other explanatory indices.

### 2.3. Econometric method

One shortcoming of Sharpe’s style analysis is the fact that it only focuses on point estimates for the factor loadings, but the main question is how much confidence we can have in the point estimates. Hence, obtaining the correct distribution and confidence intervals of the factor loadings leads us to the correct statistical arguments.

The standard OLS estimator does not suffice anymore because this usually does not lead to parameter estimates that meet the restrictions. Instead, we propose to find the factor loadings that employ with the strong form of RBSA:

$$
\begin{aligned}
&Min\;\; N^{-1}\sum_{k=1}^{N}(R_t-\beta_k I_{kt})^2 \\
&s.t.\;\; \sum \beta_k = 1\,,\quad \beta \geq 0.
\end{aligned}
\tag{5}
$$

The constraints on the factor loading, such as restrictions in (2) and (3), complicate the calculation and the problem of calculating the weights solved by using the technique of quadratic programming. Judge and Takayama (1966) show that the normal assumption for the factor loading estimator is no longer valid because the distribution is truncated at zero. Lobosco and DiBartolomeo (1997) propose a method to approximate the factor loadings estimator, but their method is not valid when one of the true style coefficients are zero or one. Andrews (1997b, 2000) introduce an alternative method for obtaining confidence intervals in constrained regressions that permits us to obtain statistically valid asymptotic measures for factor loadings. Kim et al. (2005) conclude that Andrews' method has better performance when some parameters are not at the boundary. We propose an alternative approach by using Kuhn-Tucker optimization algorithm and standard bootstrapping we obtain the asymptotic distribution of the style weights.

The asymptotic distribution plays an important role in testing the significance of factor loadings. One statistical test focuses on the ability to identify whether a fund invests in a particular type of security. Another relevant test is to check whether a fund is more invested in one type of style than in another one, also requiring the asymptotic distribution.

Equation (5) is a linear-quadratic programming problem that we estimate by applying the Kuhn-Tucker algorithm. In this section, we describe the Kuhn-Tucker algorithm steps for the Sharpe's style analysis model. For similar derivations and applications of statistical inference in constrained linear models see Gouriéroux, Holly and Monfort (1982), Gouriéroux and Monfort (1995) and Kodde and Palm (1986).

Sharpe's model as described in equation (1) – (3) is compactly rewritten in matrix algebra terms as follows:

$$Y = X\beta + u \tag{6}$$

$$j'\beta = 1 \tag{7}$$

$$\beta_k \geq 0 \qquad k = 2, \ldots, N+1 \tag{8}$$

Where Y is a $(T \times 1)$ vector of fund returns, X denotes a $T \times (N+1)$ matrix where the elements in the first column are all one, and the other columns consist of N style index returns, u is a $T \times 1$ vector of error terms. The $(N+1) \times 1$ vector $\beta$ has as first element the intercept $\alpha$ and the other elements are the style index sensitivities denoted by $\beta_k$ ($k = 1, \ldots, N$). In Sharpe's model the factor loadings add-up to one and all the factor loadings are non-negative. Note that $j$ is a $(N+1) \times 1$ vector where the first element is zero and the other elements are equal to one.

We are interested in the parameter estimates together with the associated asymptotic distribution for the vector**.** Because of the inequality constraints, we employ the estimation algorithm introduced by Kuhn-Tucker. We show that in the case of a linear regression model this Kuhn-Tucker estimator, denoted as $b_{KT}$, can be written in terms of a so-called Lagrange estimator, $b_L$. A Lagrange estimator finds optimal parameter estimates subject to equality constraints. Next, the Lagrange estimator can be expressed in terms of the ordinary least squares (OLS) unconstrained estimator, $b_U$.

The principle behind the Kuhn-Tucker algorithm lies in the treatment of the inequality constraints on the factor sensitivities. When a particular constraint is non-binding then its estimator for the associated factor loading is equal to the OLS estimator. When the particular constraint is binding then its estimator is equal to the Lagrange estimator. Beforehand it is not known which constraints will be binding and which will be non-binding. Therefore, we consider the estimators for all possible combinations of binding and non-binding restrictions. The combination that leads to the lowest residual sum of squares and that meets all constraints then leads to the optimal parameter estimates. We show that the Kuhn-Tucker solution is expressed in terms of the unconstrained least squares estimator as follows:

$$b_{KT} = \min_{S\in\Omega}\left\{\left(Y - Xb_S\right)'\left(Y - Xb_S\right) \mid j'b_S = 1; Sb_S = 0\right\} \tag{9}$$

Where

$$b_S = \left[I_{N+1} - VS'(SVS')^{-1}S\right]P + \left[I_{N+1} - VS'(SVS')^{-1}S\right]\left[I_{N+1} - P\right]b_U \tag{10}$$

$$P = \left(X'X\right)^{-1} j\left[j'\left(X'X\right)^{-1} j\right]^{-1} \tag{11}$$

$$V = \left(I_{N+1} - Pj'\right)\left(X'X\right)^{-1} \tag{12}$$

And $I_{N+1}$ is the $(N+1)\times(N+1)$ identity matrix. Let S be the matrix that represents the binding inequality constraints, i.e. the associated equality constraint reads

$$S\beta = 0 \tag{13}$$

For example, the following $2\times(N+1)$ matrix represents the sub-problem where the second and the third parameter are binding:

$$S = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \end{pmatrix} \tag{14}$$

The set of all possible matrices S representing combinations of binding and non-binding constraints is given by $\Omega$. The expressions above show that the Kuhn-Tucker solution is identical to the Lagrange estimator ($b_S$) for one of the possible sub-problems ($S \in \Omega$), i.e. combination of binding and non-binding constraints. In equation (10) we show that this estimator is related to the unconstrained estimator and some deterministic matrices. The unconstrained least squares estimator reads

$$b_U = (X'X)^{-1} X'Y \tag{15}$$

And the associated variance covariance matrix is given by

$$V(b_u) = \hat{\sigma}^2 (X'X)^{-1} \tag{16}$$

Where $\hat{\sigma}^2$ is the variance of the residuals. The asymptotic distribution of the Kuhn-Tucker estimate follows by employing the standard bootstrapping technique. To arrive at this distribution we proceed as follows:

1. Draw a sample for the error term, denoted with $u^{(i)} \sim N(0, \hat{\sigma}^2 I_T)$
2. Construct a vector of dependent variables $y^{(i)} = Xb_{KT} + u^{(i)}$
3. Estimate the model $y^{(i)} = X\beta + u^{(i)}$ subject to the constraints in (7) and (8)
4. This leads to an estimate $b_{KT}^{(i)}$
5. Repeat steps (1)-(4) 10,000 times. This gives a set $b_{KT}^{(i)} \quad i = 1, \ldots, 10{,}000$

These 10.000 values represent the asymptotic distribution of the Kuhn-Tucker estimator.

Finally, we obtain the asymptotic confidence interval by using the percentiles of the bootstraped distribution. When bootstraped samples are completed, we sort the results and then the 5$^{th}$ and 95$^{th}$ largest values show the confidence interval.

## 3. Data

One of the main limitations of previous relevant studies is that they only focus on small samples of mutual funds for their empirical tests. For example, Kim et al. (2005) use 2 U.S. mutual funds, and their sample period is from 1979 through 1997 for the Fidelity Magellan Fund and from 1991 through 1998 for the Minicap Fund. De Roon et al. (2004) use 18 U.S. based internationally mutual funds with international MSCI growth and value indices, from 1989 through 1999. Swinkels and Van der Sluis (2006) use 12 international funds and 87 asset allocation funds and finally, Kim et al. (2000) assess U.S. funds in 7 objective groups over the period from 1993 through 1996. Hence, one of the contributions in this paper is that we analyze a more extensive and thoroughly large sample of U.S. mutual funds which we retrieve from the Center for Research in Security Prices (CRSP). The CRSP mutual fund database includes information on monthly total returns[3], total net assets (TNA), different types of fees, turnover ratio, and other mutual fund characteristics.

Following many relevant studies such as Jonathan Berk and Binsbergen (2013), Andrew Mason et al. (2012), Vikram Nada et al. (2009), Marcin Kacperczyk et al. (2008), Joseph Chen et al. (2004) and Elton, Gruber and Blake (2001), we consider six selection criteria to arrive at the mutual fund data, relevant for our empirical study. Additionally, it should be noted that the total returns provided by CRSP are after fees, expenses and brokerage commissions but before front-end or back- end loads.

First, to facilitate comparison with the prior literature, we restrict our analysis to domestic U.S. equity mutual funds, so we dropped all balanced, bond, international and sector funds.

[3] Monthly total returns values are calculated as a change in NAV including reinvested dividends from one period to the next. NAVs are net of all management expenses and 12b-1 fees. Front and rear load fees are excluded.- Survive bias free U.S. mutual fund guide, October 2014

Second, because some mutual funds which have different share classes may enter into the database multiple times at the same period, we aggregate all share classes into a single fund to eliminate such redundant observations.

Third, because the CRSP mutual fund database includes some special information on fund returns which are reported voluntarily by some small mutual funds, there may be a systematic upward bias in the reported returns among these observations. Therefore, we drop all fund observations where the size of the fund in the previous quarter does not exceed $1 million.

Fourth, we check the mutual funds' asset composition and remove all funds from the database which have negative weights to exclude short-selling considerations.

Fifth, since we focus only on actively managed mutual funds and because CRSP does not provide any way to discriminate between actively and passively managed mutual funds, we remove all index funds from our sample, and we also exclude all funds which have a zero turnover ratio in one year to make sure that our sample includes actively managed funds, only.

Sixth, we include only funds that exist for at least 24 months during the estimation period.

Therefore, the number of distinct U.S. mutual funds that meet our selection criteria over the sample period from July 2003 through December 2014, is 2,637 funds. These funds are classified into four main different investment objectives based on the Lipper Prospectus Objective codes[4] which are provided by the CRSP mutual fund data base.

In addition, through 2,637 individual mutual funds, we form six value weighted portfolios, which will be analyzed in more depth. Furthermore, we calculate TNA-weighted average monthly total return. We employ an equally weighted portfolio containing all funds

[4] http://www.crsp.com/products/documentation/lipper-objective-and-classification-codes
Lipper's objective codes are assigned based on the language that the fund uses in its prospectus to describe how it intends to invest.

individually, an equally weighted portfolio of funds that did not survive during the whole sample period (dead funds), and value weighted portfolios of mutual funds within a particular investment objective: growth/income, income, growth and small cap.

As we only consider U.S. equity funds, the relevant style benchmarks are all U.S. indices which are all monthly total returns. We include the U.S. value index (S&P500 Value index), the U.S. growth index (S&P500 Growth index), the U.S. small cap index (S&P600 index) and two fixed income classes, cash (30-day Treasury bill rate) and bonds (30 years bonds). We retrieve the indices data from the FactSet Research System Inc. and obtain fixed income data from CRSP. This results in a 5-factor model, which is used to determine a fund's style. Summary statistics on the different equally weighted portfolios (panel A) and benchmarks (panel B) are provided in table 1.

In the RBSA method, it is important to investigate the impact of multicollinearity between different benchmarks. As De Roon et al. (2004) discuss, the difference between actual portfolio holdings in mutual funds and estimated exposures by using RBSA is more likely to be caused by the correlations between the different indices, related to multicollinearity problems. Hence, to figure out the multicollinearity among benchmarks, we employ Pearson's correlation coefficients and Variance Inflation Factors (VIF) which have been introduced in the previous section. Table 1 in panel C shows that all VIFs are less than 10 which means multicollinearity among benchmarks is not too much. But because the correlation between them is relatively high, we employ Russell indices in the robustness test section to investigate the sensitivity of our results to benchmarks.

Furthermore, it appears that funds focusing on smaller companies deliver the highest performance (11.63%) during the sample period. This however is also associated with the highest standard deviation (19.4%). Note that over 35% of the funds in our sample did not survive during the entire period. Because these funds under-perform the average fund by about

2%, A severe survivorship bias could arise if they were excluded. In our subsequent analysis, we therefore consider them as “Dead” portfolio.

In addition, table 2 reports summary statistics on fund total net assets (TNA), 52 week low Net Asset Value (NAV), 52 week high Net Asset Value (NAV), age, expense ratio, turnover ratio, management fee, and institutional fund proportion using the CRSP mutual fund data.

***Table 1***
**Summary statistics, 2003- 2014**

**Panel A:** Mutual fund returns

| Investment objective | Mean Return | Standard Deviation | Number of funds |
|---|---|---|---|
| Growth/Income | 8.81 | 14.5 | 494 |
| Income | 8.92 | 14.7 | 111 |
| Growth | 10.01 | 16.3 | 704 |
| Small caps | 11.63 | 19.4 | 337 |
| All funds | 9.67 | 14.62 | 1646 |
| Dead funds | 7.73 | 14.83 | 990 |

**Panel B:** Benchmark returns

| | | | Cross correlations | | | | |
|---|---|---|---|---|---|---|---|
| Benchmark | Mean Return | Standard deviation | Value | Growth | Small cap | Cash | Bond |
| S&P 500 Value | 8.33 | 15.48 | 1.00 | | | | |
| S&P 500 Growth | 8.71 | 13.61 | 0.92 | 1.00 | | | |
| S&P 600 Small cap | 11.08 | 18.66 | 0.89 | 0.87 | 1.00 | | |
| 30 Day treasury Bill | 1.36 | 0.51 | -0.03 | -0.07 | -0.06 | 1.00 | |
| 30 Years Bonds | 7.67 | 14.75 | -0.29 | -0.31 | -0.34 | -0.01 | 1.00 |

*Notes* This table provides summary statistics on the U.S. mutual funds (Panel A) and benchmarks (Panel B) that are used to perform the Sharpe asset class factor model. Panel A reports annualized total returns with corresponding standard deviations for six equally weighted portfolios of funds. Panel B reports returns and standard deviations on the benchmarks that are used. Finally, cross correlations between the benchmarks are given in Panel B columns 4 through 8.

**Panel C**: Variance inflation factors in the proposed benchmarks

| | S&P 500 Value | S&P 500 Growth | S&P 600 Small cap | 30 Day treasury Bill | 30 Years Bonds |
|---|---|---|---|---|---|
| $R^2$ | 0.89 | 0.85 | 0.83 | 0.02 | 0.12 |
| VIF | 9.09 | 6.66 | 5.88 | 1.02 | 1.13 |

*Notes VIF (*variance inflation factors defines) the degree to which each of the five benchmarks contributes to the multicollinearity of the model. This table reports the VIF in the proposed benchmarks and $R^2$ is the determinant coefficient in a linear regression of the explanatory index i in relation to the other four explanatory indices.

**Table 2**
**Summary statistics of mutual fund characteristics, 2003- 2014**

| Variable | Mean | Std. Dev. | Median |
|---|---|---|---|
| Total Net Assets (TNA) (in Millions) | 154.16 | 570.94 | 35.2 |
| 52 week low Net Asset Value (NAV) (in Millions) | 16.17 | 14.54 | 12.75 |
| 52 week high Net Asset Value (NAV) (in Millions) | 19.93 | 17.74 | 15.63 |
| Age (in Years) | 8.56 | 2.50 | 8.37 |
| Expense Ratio (in Percent) | 1.26 | 0.54 | 1.20 |
| Turnover Ratio | 0.64 | 0.62 | 0.49 |
| Management Fee | 0.43 | 1.25 | 0.64 |
| Institutional Fund Proportion | 0.37 | | |
| Total Number of Funds | 2,637 | | |
| Total Number of Observations | 181,852 | | |

This table provides summary statistics of characteristics of the mutual funds in our sample between July 2003 and December 2014. Fund Turnover Ratio is Minimum (of aggregated sales or aggregated purchases of securities), divided by the average 12-month Total Net Assets of the fund. Management Fee which means its fee ($) divided by Average Net Assets ($). The management fee is calculated using ratios based on the line items reported in the Statement of Operations. The management fee can be offset by fee waivers and/or reimbursements which will make this value differ from the contractual fees found in the prospectus. The proportion of Institutional mutual funds among the whole sample in percentage.

## 4. Results

An explicit change of the stated investment objective which is written in the funds' prospectus requires approval from the shareholders and it also may force some existing investors to close their accounts[5]. Kim et al. (2000) found that more than 92% of mutual funds did not change their stated objective over their sample period. For this reason, in our analysis, we assume funds' investment objective to be constant over time.

We first report the result of statistical tests for return based style analysis by using the bootstrapping method and clarify the main mechanisms through which mutual funds deviate from their benchmark. Hence, we are able to separate all mutual funds into two groups henceforth referred to the well-classified and misclassified funds. Therefore in section 4.1, we focus on the question whether serious style deviations by U.S. mutual fund managers (still) exist. Then, in section 4.2, we discuss the performance consequences of this deviation behavior to understand its impact on investors.

Using various methods of fund performance evaluation, we compare the performance of well-classified and misclassified mutual funds. We find that mutual funds which deviate from their benchmark tend to subsequently perform worse than funds that stick to their benchmark. Hence, there is a negative relation between objective misclassification and associated mutual fund performance.

The performance consequences are particularly severe for funds that significantly deviate from their benchmark. Moreover, we investigate potential reasons for misclassified mutual fund's poor performance. Finally in section 4.3 we seek to explain fund managers' misclassification and suggest that misclassification is caused by agency issues.

---

[5] Section 13 (a), item (3), of investment Company Act of 1940, version January 3, 2012, states that "deviation from its policy in respect of concentration of investments in any particular industry or group of industries as recited in its registration statement, deviation from any investment policy which is changeable only if authorized by shareholder vote".

## 4.1- Statistical Tests for Return-Based Style Analysis

In the empirical application we estimate Sharpe's model for the six value weighted portfolios that have been introduced in the previous section. Second, we determine the asymptotic distribution for the style weights by using bootstrapping method. In our analysis we focus on the added value of the extra statistical information available in the asymptotic distribution of the parameter estimates.

This distribution is applied to perform a series of tests. We first concentrate our efforts on the preciseness and significance of the style weights (table 3). Then, we check whether specific factor loadings are significantly different from each other (table 4). Finally, we test for misclassification (table 5). This answers the question whether a fund with a particular objective actually is for the largest part invested in the correct style.

To emphasize the difference between funds that ceased to exist and those that are alive, we provide results both for "Dead funds" and funds that remain "Alive".

Table 3 reports the parameter estimates of Sharpe's model for six different value weighted portfolios. In panel A estimated style weights are given. Each row deals with one particular investment objective, where the elements in columns 2 to 6 report the estimated style weights.

Interpreting the estimated weights as an approximation of portfolio holdings makes it possible to check whether funds adhere to their stated investment objective. We find that the income fund/portfolio is mainly exposed to the value benchmark, the growth fund/portfolio to the growth benchmark and finally the small cap fund/portfolio is up to 81% exposed to the smaller companies benchmark both for Alive and Dead funds. Our results suggest that mutual funds invest as they are supposed to do.

In panel B 95% confidence intervals are given for all factor loadings. The confidence intervals show that the point estimates are relatively precise reflections of the portfolio weights.

**Table 3**
**Results Sharpe asset class factor model**

**Panel A:** Estimated style weights for Alive Portfolio

| Objective | Value | Growth | Small cap | Cash | Bond | $R^2$ |
|---|---|---|---|---|---|---|
| Growth | 0.15*** | 0.63*** | 0.21*** | 0.00 | 0.00 | 0.98 |
| Income | 0.53*** | 0.35*** | 0.03*** | 0.07*** | 0.01*** | 0.98 |
| Small cap | 0.00 | 0.18*** | 0.82*** | 0.00 | 0.00 | 0.98 |
| Growth/Income | 0.35*** | 0.39*** | 0.14*** | 0.10*** | 0.01*** | 0.92 |

**Panel B:** 95% Confidence intervals for style weights for Alive Portfolio

| Objective | Value | Growth | Small cap | Cash | Bond |
|---|---|---|---|---|---|
| Growth | [0.07 – 0.24] | [0.55 – 0.71] | [0.16 – 0.27] | [0.00 – 0.05] | [0.00 – 0.00] |
| Income | [0.48 – 0.59] | [0.30 – 0.40] | [0.00 – 0.03] | [0.05 – 0.11] | [0.00 – 0.01] |
| Small cap | [0.00 – 0.00] | [0.14 – 0.34] | [0.76 – 0.87] | [0.00 – 0.00] | [0.00 – 0.00] |
| Growth/Income | [0.26 – 0.43] | [0.24 – 0.54] | [0.03 – 0.25] | [0.07 – 0.15] | [0.00 – 0.02] |

**Panel C:** Estimated style weights for Dead Portfolio

| Objective | Value | Growth | Small cap | Cash | Bond | $R^2$ |
|---|---|---|---|---|---|---|
| Growth | 0.09*** | 0.64*** | 0.26*** | 0.00 | 0.00 | 0.96 |
| Income | 0.50*** | 0.27*** | 0.00*** | 0.20*** | 0.02*** | 0.96 |
| Small cap | 0.00 | 0.19*** | 0.81*** | 0.00 | 0.00 | 0.97 |
| Growth/Income | 0.43*** | 0.41*** | 0.08*** | 0.07*** | 0.00 | 0.98 |

**Panel D:** 95% Confidence intervals for style weights for Dead Portfolio

| Objective | Value | Growth | Small cap | Cash | Bond |
|---|---|---|---|---|---|
| Growth | [0.00 – 0.20] | [0.54 – 0.73] | [0.16 – 0.37] | [0.00 – 0.05] | [0.00 – 0.00] |
| Income | [0.44 – 0.57] | [0.22 – 0.33] | [0.00 – 0.00] | [0.18 – 0.25] | [0.00 – 0.02] |
| Small cap | [0.00 – 0.00] | [0.15 – 0.35] | [0.75 – 0.87] | [0.00 – 0.00] | [0.00 – 0.00] |
| Growth/Income | [0.36 – 0.50] | [0.36 – 0.47] | [0.03 – 0.13] | [0.05 – 0.11] | [0.00 – 0.00] |

*Notes* This table presents the parameter estimates of the Sharpe return-based model for six value weighted portfolios of funds. In panel A estimated style weights are given. Each row deals with one particular investment objective, where the elements in columns 3 to 7 report the estimated style weights. Panel B reports the 95% confidence intervals for all estimated style weights. Because of the constraints on the parameters these have been constructed by bootstrapping.

*** Significantly different from zero at the 1 % level
** Significantly different from zero at the 5 % level
* Significantly different from zero at the 10% level
With Significance in Panel A based on confidence intervals reported in Panel B.

In table 4, we deal with the question whether two style weights are significantly different from each other. For each value weighted portfolio we compare all five factor loadings with each other. This leads to 10 comparisons per value weighted portfolio. We examine whether the difference between two style weights is significantly different from zero. The distribution of the differences follows directly from bootstrapping. Using a simple p-test we then determine whether this difference is statistically significant.

Given are the mean difference (column 4) and whether the mean is significantly different from zero at the 5% level (column 5). From the results it appears that exposures towards the T-bill (cash) and Bond index often are not significantly different from each other. If we turn back to our earlier observations on style adherence it still seems that most funds closely follow their style. For instance, income funds have a value weight that is significantly larger than their growth and/or small cap weight.

**Table 4**
**Significance for differences between estimated Sharpe style weights**

**Panel A: Alive mutual funds**

| Objective | index1 | index2 | Mean difference | Significant? |
|---|---|---|---|---|
| Growth | Value | Growth | -0.48 | YES |
| Growth | Value | Small cap | -0.06 | NO |
| Growth | Value | Bond | 0.15 | YES |
| Growth | Value | T-Bill | 0.14 | YES |
| Growth | Growth | Small cap | 0.42 | YES |
| Growth | Growth | Bond | 0.63 | YES |
| Growth | Growth | T-Bill | 0.62 | YES |
| Growth | Small cap | Bond | 0.21 | YES |
| Growth | Small cap | T-Bill | 0.20 | YES |
| Growth | Bond | T-Bill | -0.01 | NO |
| Income | Value | Growth | 0.18 | YES |
| Income | Value | Small cap | 0.51 | YES |
| Income | Value | Bond | 0.52 | YES |
| Income | Value | T-Bill | 0.45 | YES |
| Income | Growth | Small cap | 0.33 | YES |
| Income | Growth | Bond | 0.34 | YES |
| Income | Growth | T-Bill | 0.27 | YES |
| Income | Small cap | Bond | 0.01 | NO |
| Income | Small cap | T-Bill | -0.06 | YES |
| Income | Bond | T-Bill | -0.07 | YES |
| Small cap | Value | Growth | -0.22 | YES |
| Small cap | Value | Small cap | -0.81 | YES |
| Small cap | Value | Bond | 0.00 | YES |
| Small cap | Value | T-Bill | 0.00 | NO |
| Small cap | Growth | Small cap | -0.59 | YES |
| Small cap | Growth | Bond | 0.22 | YES |
| Small cap | Growth | T-Bill | 0.22 | YES |
| Small cap | Small cap | Bond | 0.81 | YES |
| Small cap | Small cap | T-Bill | 0.81 | YES |
| Small cap | Bond | T-Bill | 0.02 | NO |
| Growth/Income | Value | Growth | -0.11 | YES |
| Growth/Income | Value | Small cap | 0.26 | YES |
| Growth/Income | Value | Bond | 0.35 | YES |
| Growth/Income | Value | T-Bill | 0.23 | YES |
| Growth/Income | Growth | Small cap | 0.37 | YES |
| Growth/Income | Growth | Bond | 0.46 | YES |
| Growth/Income | Growth | T-Bill | 0.34 | YES |
| Growth/Income | Small cap | Bond | 0.09 | YES |
| Growth/Income | Small cap | T-Bill | -0.03 | NO |
| Growth/Income | Bond | T-Bill | -0.11 | YES |

**Panel B: Dead mutual funds**

| Objective | index1 | index2 | Mean difference | Significant? |
|---|---|---|---|---|
| Growth | Value | Growth | -0.54 | YES |
| Growth | Value | Small cap | -0.17 | NO |
| Growth | Value | Bond | 0.10 | NO |
| Growth | Value | T-Bill | 0.09 | NO |
| Growth | Growth | Small cap | 0.37 | YES |
| Growth | Growth | Bond | 0.64 | YES |
| Growth | Growth | T-Bill | 0.63 | YES |
| Growth | Small cap | Bond | 0.27 | YES |
| Growth | Small cap | T-Bill | 0.26 | YES |
| Growth | Bond | T-Bill | -0.02 | NO |
| Income | Value | Growth | 0.23 | YES |
| Income | Value | Small cap | 0.52 | YES |
| Income | Value | Bond | 0.49 | YES |
| Income | Value | T-Bill | 0.31 | YES |
| Income | Growth | Small cap | 0.29 | YES |
| Income | Growth | Bond | 0.26 | YES |
| Income | Growth | T-Bill | 0.08 | YES |
| Income | Small cap | Bond | -0.03 | YES |
| Income | Small cap | T-Bill | -0.21 | YES |
| Income | Bond | T-Bill | -0.18 | NO |
| Small cap | Value | Growth | -0.25 | YES |
| Small cap | Value | Small cap | -0.83 | YES |
| Small cap | Value | Bond | -0.02 | NO |
| Small cap | Value | T-Bill | -0.02 | YES |
| Small cap | Growth | Small cap | -0.58 | YES |
| Small cap | Growth | Bond | 0.23 | YES |
| Small cap | Growth | T-Bill | 0.23 | YES |
| Small cap | Small cap | Bond | 0.81 | YES |
| Small cap | Small cap | T-Bill | 0.82 | YES |
| Small cap | Bond | T-Bill | 0.00 | NO |
| Growth/Income | Value | Growth | 0.01 | NO |
| Growth/Income | Value | Small cap | 0.35 | YES |
| Growth/Income | Value | Bond | 0.44 | YES |
| Growth/Income | Value | T-Bill | 0.35 | YES |
| Growth/Income | Growth | Small cap | 0.34 | YES |
| Growth/Income | Growth | Bond | 0.43 | YES |
| Growth/Income | Growth | T-Bill | 0.34 | YES |
| Growth/Income | Small cap | Bond | 0.09 | YES |
| Growth/Income | Small cap | T-Bill | 0.00 | YES |
| Growth/Income | Bond | T-Bill | -0.08 | NO |

*Notes* This table addresses the question whether two style weights are significantly different from each other by examining whether the *difference* between two weights is significantly different from zero. The distribution of the differences follows directly from the results of our bootstrap. Given are the mean differences (column 4) and whether the mean is significantly different from zero at the 5% level (column 5). For each of our value weighted portfolios of funds we compare all five weights with each other. This leads to 10 comparisons per value weighted portfolio.

The analysis on the value weighted portfolios as provided in table 3 and 4 does not produce evidence of serious style deviations by mutual fund managers. It may be the case that the construction of value weighted portfolios averages out effects that are present in individual funds. Next, we therefore consider the mutual fund misclassification phenomenon by analyzing the historical returns of all *individual* mutual funds separately.

We assume that a growth/income fund should predominantly be exposed to the growth or value benchmark, income funds to the value benchmark, growth funds to the growth benchmark and finally small cap funds to the small cap benchmark. If a fund exhibits a higher weight on any other benchmark, we consider it to be misclassified. In table 5 we summarize the results of this exercise. In column 3 we also take into account the information in the asymptotic distribution function. On average 29% of all funds is predominantly exposed to a benchmark other than the one we would expect it to be exposed to. Especially growth funds tend to be misclassified whereas small cap funds adhere to their style for 100%. These results are in line with prior studies such as Mason et al. (2012) who find that for U.S. mutual funds for the periods 1996 – 1998 and 2003 – 2005, the large cap growth and small cap value funds are significantly exposed to other benchmarks. Kim et al. (2005) also find that the Fidelity Magellan Funds with large growth style is not only oriented towards large cap growth but also is exposed to the large cap value. In addition, DiBartolomeo & Witkowski (1997), Indro et al (1998) and Kim, Shukla & Tomas (2000) find that more than 50% of U.S. mutual funds differ from their benchmarks, and over 30% of the funds are severely misclassified. These observations are based on point estimates of style weights.

The result of this exercise is summarized in column 3. If we take into account the significance of estimated style weights we find that 25% of the mutual funds differ from their stated investment objective and over 18% of the funds are severely misclassified. Moreover, 25% of dead funds are persistently misclassified.

**Table 5**

**Mutual fund misclassifications**

**Panel A:** Mutual fund misclassifications based on individual fund returns (Alive)

| Objective | % Misclassifications | % Significant Misclassifications |
|---|---|---|
| Growth/Income | 32% | 23% |
| Income | 31% | 18% |
| Growth | 29% | 17% |
| Small Cap | 3% | 0% |

**Panel B:** Mutual fund misclassifications based on individual fund returns (Dead)

| Objective | % Misclassifications | % Significant Misclassifications |
|---|---|---|
| Growth/Income | 28% | 18% |
| Income | 38% | 29% |
| Growth | 33% | 15% |
| Small Cap | 5.5% | 3.3% |

*Notes* This table presents evidence of fund misclassification using individual fund returns. We assume that a growth/income fund should predominantly be exposed to the growth or value benchmark, income finds to the value benchmark, growth funds to the growth benchmark and finally small cap funds to the small cap benchmark. If a fund exhibits a higher weight on any other benchmark, we consider it to be misclassified. Column 2 reports the percentage of misclassified funds per investment objective, solely based on the point estimates for style weights. In column 3 we take into account the significance of estimated style weights and report the percentage of significantly misclassified funds per investment objective.

## 4.2- Objective gaming and Performance Consequences of misclassification

The previous section suggests that some mutual fund managers change their investment objective style over time. Possible reasons to deviate from a stated investment style is objective gaming which has gained widespread attention among investment professionals. Using this strategy, the fund managers may be expected to take advantage of their market timing ability to perform better in comparison to the funds in his stated objective group. Hence, we compare the mean difference return between misclassified and well-classified mutual funds to investigate whether misclassified funds that deviate from their stated investment objective out-perform relative to their peers.

A difference in return between misclassified and well-classified funds is reported per investment objective in table 6. The evidence shows that on average the annual return of misclassified funds is 0.86% lower than the return on well-classified funds. If we take the significance of estimated style weights into account, leading to a different set of misclassified funds, panel B arises. Again it appears that misclassified funds under-perform well-classified funds and on average the annual return of misclassified funds is 1.08% lower than the return on well-classified funds.

**Table 6**

**Objective Gaming**

**Panel A:** Without Confidence intervals

| Objective | Return Misclassified | Return Well-classified | Difference (Miss-Well) |
|---|---|---|---|
| Growth/Income | 7.87 | 8.61 | -0.75 |
| Income | 9.02 | 9.29 | -0.27 |
| Growth | 9.37 | 9.38 | -0.01 |
| Small Cap | 10.98 | 11.73 | -0.74 |
| All | 9.01 | 9.88 | -0.86 |
| Dead | 7.54 | 7.78 | -0.24 |

**Panel B:** With Confidence intervals

| Objective | Return Misclassified | Return Well-classified | Difference (Miss-Well) |
|---|---|---|---|
| Growth/Income | 7.65 | 8.58 | -0.94 |
| Income | 8.53 | 9.35 | -0.83 |
| Growth | 9.37 | 9.48 | 0.10 |
| Small Cap | -- | 15.65 | -- |
| All | 8.74 | 9.82 | -1.08 |
| Dead | 7.65 | 7.68 | -0.03 |

*Notes* In this table we examine whether funds that are misclassified out-perform their peers, in other words does objective gaming pay? As before, we assume that a growth/income fund should predominantly be exposed to the growth or value benchmark, income funds to the value benchmark, growth funds to the growth benchmark and finally small cap funds to the small cap benchmark. If a fund exhibits a higher weight on any other benchmark, we consider it to be misclassified. In Panel A column 4 the (annualized) difference in return between misclassified and well-classified funds is reported per investment objective. Note that here only the point estimates of style weights are considered in detecting misclassified funds. Panel B takes the significance of estimated style weights into account when forming the misclassified group and then similarly reports annual return differences in column 4.

To gain more insight into objective gaming behavior, we examine the risk adjusted return of misclassified funds to assess the potential risks that investors are exposed to.

Chevalier and Ellison (1997), Sirri and Tufano (1998), and Huang et al. (2007) find a convex relation between fund flows and their performance which means investors tend to invest in funds with a great performance and do not penalize poor performance equivalently. This convex flow-performance relation can motivate fund managers to attract more investors to strategically invest in the "Hot" style or away from the current "Cold" style regardless of their stated investment objective. As mentioned before, as long as this deviation leads fund managers to perform better than other funds, it benefits investors, otherwise they can only impose higher risk to investors. Huang, Sialm & Zhang (2011) (also) find that mutual funds that increase risk underperform funds that keep stable risk levels over time. They argue that risk shifting between mutual funds is motivated by agency issues. Hence, to take into account the risk effect on fund performance, we examine the abnormal returns of funds by computing the excess return relative to the value-weighted market portfolio. We employ the one-factor CAPM, the Fama and French (1993), the Carhart (1997), and the Pastor and Stambaugh (2003) models. The Carhart model is specified as follows:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_{1,i}MKT_t + \beta_{2,i}SMB_t + \beta_{3,i}HML_t + \beta_{4,i}MOM_t + \varepsilon_{i,t}$$

where, the TNA- weighted average return of fund *i* during time period *t* is denoted by$R_{i,t}$. Also, MKT is excess market return, SMB is small-minus-big, HML is high-minus-low and MOM is the momentum factor to the risk-adjusting model to control for returns on momentum trading strategies. MKT, SMB, HML, MOM are taken from Kenneth French's website. The Carhart model nests the CAPM model (which includes only the market factor) and the Fama-French model (which includes the size and the book-to market factors in addition to the market factor).

The traded liquidity factor has been defined by Pastor and Stambaugh (2003) as the value-weighted return on the 10-1 portfolio from a sort of stocks into decile groups depending on their historical liquidity betas, or stock sensitivities to innovations in the aggregate liquidity.

The Pastor-Stambaugh model adds a liquidity factor as a fifth factor to the Carhart model. Hence, the 5-factor model is specified as follows:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_{1,i}MKT_t + \beta_{2,i}SMB_t + \beta_{3,i}HML_t + \beta_{4,i}MOM_t + \beta_{5,i}LIQ_t + \varepsilon_{i,t}$$

Where LIQ is a traded liquidity factor. Other factors are the same as in the Carhart model. Traded liquidity data is obtained through Wharton Research Data Services.

We form portfolios of misclassified and well-classified funds and estimate the risk exposures based on the return of these portfolios. We calculate the fund portfolio return as the equally-weighted mean return of all funds over the sample period. Table 7 reports the result based on the mutual fund investment objective classification. Panel A and C in this table report the excess return of well-classified and misclassified funds respectively. Panel B and D show the excess return of misclassified and well-classified funds when we take the significant of estimated style weights into account. We observe that funds which deviate from their benchmark have worse subsequent risk-adjusted returns than funds which stick to their benchmark, especially when we consider the significance of estimated style weights. The poor performance of misclassified funds remains economically and statistically significant using alternative factor models. For example, using the Fama- French model, we find that misclassified funds on average have an abnormal return around -6.8 basis points per month, while this number is -4.8 basis points for well-classified mutual funds. In all four models, the factor loadings denote the sensitivities of the returns of portfolio *i* to the various factors. As documented in table 7, results are robust and all models proof that misclassified funds under perform funds that stick to their benchmarks. In addition, we observe that the performance of Dead funds in all models is significantly negative.

In addition, these results are in line with De Roon et al (2004) who consider U.S. based internationally mutual funds and find that the actively managed funds in Foreign and World styles relatively underperform their corresponding mimicking portfolio that is a combination of the MSCI indices.

**Table 7**

**Performance consequence of investment objective deviation for well-classified funds**

**Panel A:** Well-classified funds without confidence intervals

| **Objective** | **CAPM** | **Fama- French** | **Carhart** | **Pastor-Stambaugh** |
|---|---|---|---|---|
| Growth/Income | -0.084*** | -0.082*** | -0.077*** | -0.083*** |
| Income | -0.014 | -0.014 | -0.008 | -0.003 |
| Growth | -0.049 | -0.045 | -0.05 | -0.064* |
| Small Cap | -0.006 | -0.036 | -0.036 | -0.052 |
| All | -0.042 | -0.048* | -0.049* | -0.061* |
| Dead | -0.223*** | -0.228*** | -0.227*** | -0.225*** |

**Panel B:** Well-classified funds with confidence intervals

| **Objective** | **CAPM** | **Fama- French** | **Carhart** | **Pastor-Stambaugh** |
|---|---|---|---|---|
| Growth/Income | -0.079*** | -0.076*** | -0.071*** | -0.078*** |
| Income | -0.002 | -0.001 | 0.001 | 0.0035 |
| Growth | -0.053 | -0.051 | -0. 05 | -0. 068* |
| Small Cap | -0.006 | -0.035 | -0.035 | -0.051 |
| All | -0.043 | -0.048* | -0.049* | -0.061* |
| Dead | -0.223*** | -0.228*** | -0.228*** | -0.224*** |

Notes this table reports the intercepts from factor regressions based on the CAPM, Fama-French, Carhart, and Pastor-Stambaugh models. All returns are expressed in % per month. The significance levels are abbreviated with asterisks: *, **, and *** denote significance at the 15%, 10%, and 5% levels, respectively.

**Table 7**

**Performance consequence of investment objective deviation for misclassified funds**

**Panel C:** Misclassified funds without confidence intervals

| **Objective** | **CAPM** | **Fama- French** | **Carhart** | **Pastor-Stambaugh** |
|---|---|---|---|---|
| Growth/Income | -0.038 | -0.038 | -0.033 | -0.047 |
| Income | 0.002 | 0.007 | 0.000 | -0.023 |
| Growth | -0.094** | -0. 10*** | -0.086*** | -0.099*** |
| Small Cap | 0.002 | -0.002 | 0.000 | 0.000 |
| All | -0.056 | -0.06** | -0.051 | -0.064** |
| Dead | -0.188*** | -0.18*** | -0.19*** | -0.21*** |

**Panel D:** Misclassified funds with confidence intervals

| **Objective** | **CAPM** | **Fama- French** | **Carhart** | **Pastor-Stambaugh** |
|---|---|---|---|---|
| Growth/Income | -0.044 | -0.044 | -0.038 | -0.054 |
| Income | -0.042 | -0.035 | -0.042 | -0.069 |
| Growth | -0.096** | -0. 11*** | -0.086*** | -0.096*** |
| Small Cap | - | - | - | - |
| All | -0.064** | -0.068** | -0.056** | -0.071*** |
| Dead | -0.162*** | -0.162*** | -0.163*** | -0.191*** |

Notes this table reports α from factor regressions based on the CAPM, Fama-French, Carhart, and Pastor-Stambaugh models. All returns are expressed in % per month. The significance levels are abbreviated with asterisks: *, **, and *** denote significance at the 15%, 10%, and 5% levels, respectively.

Table 8 also shows that the mean difference of alpha between misclassified and well-classified funds is always greater than zero which means well-classified funds have a higher alpha. Although the Carhart and Pastor models show that the poor performance of misclassified funds is statistically significant, results do not seem economically significant. For example, the annualized mean difference between funds is 10.8 basis point using the Pastor-Stambaugh model and 8.4 basis point when we employ Carhart model by taking the significance of estimated style weights into account. According to the results of table 7 and 8, we can argue that there is negative relation between objective gaming and fund performance. In sum, on both an unadjusted and risk-adjusted basis, misclassified funds are below- average performers.

**Table 8**

**Performance consequence of investment objective deviation for all mutual funds**

Mean difference between well-classified and misclassified funds

| Objective | CAPM | Fama- French | Carhart | Pastor-Stambaugh |
|---|---|---|---|---|
| All funds[without CI] | 0.014 | 0.012 | 0.002*** | 0.003*** |
| All funds [with CI] | 0.021 | 0.019 | 0.007** | 0.009* |

*Notes* mean difference between well-classified and misclassified funds based on four different alphas. The significance levels are abbreviated with asterisks: *, **, *** and **** denote significance at the 20%, 15%, 10%, and 5% levels, respectively. CI in bracket means confidence interval.

## 4.3- Characteristics of misclassified funds

To clarify the misclassification behaviour, we first try to answer this question what kinds of mutual funds are more likely to deviate from their stated investment objective. Hence, we investigate the misclassified mutual funds specifications from different approaches such as turnover ratio, fund size, expense ratio and institutional or retail funds.

As Carhart (1997), Cremers & Pareek (2014) and Huang, Sialm & Zhang (2011) discussed, mutual fund performance is negatively affected by the amount of trading. Hence, misclassified mutual fund's poor performance might be caused by the additional trading costs. Since these trading costs are deducted from a mutual fund's gross income, investors suffer from these costs and they may be interested in a potentially cheaper alternative. As Chalmers, Edelen and Kadlec (1999) describe, the turnover ratio captures a substantial fraction of the trading costs. Therefore, to compare the amount of trading costs of well-classified and misclassified funds, we employ the turnover ratio as a proxy for trading cost.

Table 9 shows that the misclassified mutual funds tend to exhibit a higher turnover ratio than well-classified funds. We also find that the mean difference between turnover ratio of misclassified funds and well-classified funds for Growth/Income and Income style is significant at the 5% level.

For example, the turnover ratio difference between misclassified and well-classified funds in Growth/Income style is 5.18 per month and this difference for Income style is 16.96. Therefore, misclassified mutual funds have higher trading cost. Thus trading costs can explain the poor performance of misclassified funds.

**Table 9**

**Turnover ratio as trading cost for funds with confidence interval**

| Objective | Turnover ratio Misclassified | Turnover ratio Well-classified | Difference (Miss-Well) |
|---|---|---|---|
| Growth/Income | 43.74 | 38.55 | 5.18*** |
| Income | 58.88 | 41.92 | 16.96*** |
| Growth | 52.41 | 52.30 | 0.09 |
| Small Cap | -- | 69.56 | 0.00 |
| All | 49.51 | 47.39 | 2.12 |

*Notes* this table reports the turnover ratio of misclassified and well-classified funds for each style. Significance levels are abbreviated with asterisks: *, **, *** denote significance at the 15%, 10%, and 5% levels, respectively.

Odean (1998 and 1999) argue that one explanation for unnecessary turnover is overconfidence bias which means there is direct link between investor overconfidence and excessive trading. Another explanation is proposed by Dow and Gorty (1997) who find a misalignment of interest between asset managers and asset owners. He propose a principal-agent model in which explain that portfolio managers, could rationally trade highly to send a positive signal to their investors and superiors who cannot distinguish between ‘actively doing nothing’ and ‘simply doing nothing’ and to show that they do not have passive management strategy. By increasing in delegated assets under management, the agency problem between asset owners and asset managers will be arise.

Furthermore, Gil-Bazo and Ruiz-Verdu (2009) argue that there is a negative relation between fees and performance which means high-expense funds do not perform better than low-expense funds, even before subtracting expenses. They interpret this evidence as an agency problem in which high-expense funds target less performance- sensitive investors who are naive investors and they are not responsive to expenses, hence these funds are able to charge them higher fees. Thus, high-expense funds may have bigger incentives to deviate from their stated investment objective to reach better performance and attract more fund flows. We therefore consider the annual expense ratios which are computed based on the equally- weighted average return across the share classes, and consistent with this hypothesis, we find that high expense funds are more likely to deviate. Although in general, there is no evidence showing a significant difference between expenses ratio of all misclassified and well-classified funds, table 10 reports that at least for Growth and Growth/income investment styles, the difference is statistically significant. The fact that these characteristics are related to the potential benefits of deviation provides supportive evidence that agency problems contribute to misclassification behavior.

**Table 10**
**Annual expense ratio effect on the investment objective deviation**

**Panel A: Expense ratio-without Confidence interval**

| Objective | Expense ratio Misclassified | Expense ratio Well-classified | Difference (Miss-Well) |
|---|---|---|---|
| Growth/Income | 0.77 | 0.70 | 0.07* [0.15] |
| Income | 1.08 | 1.04 | 0.04 [0.49] |
| Growth | 1.02 | 0.95 | 0.06 [0.20] |
| Small Cap | - | 1.19 | - |
| All | 0.93 | 0.88 | 0.05 [0.29] |

*Notes* this table reports the annual expense ratio of misclassified and well-classified funds for each style. Significance levels are abbreviated with asterisks: *, **, *** denote significance at the 15%, 10%, and 5% levels, respectively. The numbers in square brackets are the p-value of coefficients.

**Panel B: Expense ratio-with Confidence interval**

| Objective | Expense ratio Misclassified | Expense ratio Well-classified | Difference (Miss-Well) |
|---|---|---|---|
| Growth/Income | 0.77 | 0.70 | 0.07* |
| Income | 1.11 | 1.03 | 0.08* |
| Growth | 0.96 | 0.96 | 0.00 |
| Small Cap | - | 1.18 | - |
| All | 0.90 | 0.88 | 0.02 |

*Notes* this table reports the annual expense ratio of misclassified and well-classified funds for each style. Significance levels are abbreviated with asterisks: *, **, *** denote significance at the 15%, 10%, and 5% levels, respectively. The numbers in square brackets are the p-value of coefficients.

Huang, Wei, and Yan (2007) find that funds in smaller size have a more convex flow-performance relation. As table 11 shows, we measure fund size based on their TNA in the previous month. Hence, we observe that the misclassified mutual funds are relatively greater in terms of TNA than well-classified funds. For example, the misclassified funds for the Growth/Income and Income investment styles are significantly greater than well-classified funds. Consistent with the popular notion, small funds are indeed more active than large funds, thus in small funds, fund managers have a higher incentive to deviate from their benchmarks.

**Table 11**
**Fund size effect on the investment objective deviation**

**Panel A: fund size for funds with confidence interval**

| **Objective** | **Size score Misclassified** | **Size score Well-classified** | **Difference (Miss-Well)** |
|---|---|---|---|
| Growth/Income | 290.67 | 890.77 | -600.10*** |
| Income | 155.07 | 510.20 | -355.13*** |
| Growth | 657.65 | 404.95 | 252.70 |
| Small Cap | | | |
| All | 447.62 | 469.32 | -21.70 |

*Notes* this table reports the fund size of misclassified and well-classified funds for each style. Significance levels are abbreviated with asterisks: *, **, *** denote significance at the 15%, 10%, and 5% levels, respectively.

We also find that more than 35% of well-classified funds are institutional funds while this number for misclassified funds is 26%. This evidence shows that generally institutional funds stick more to their benchmarks than retail funds.

## 5. Robustness tests

In this section we employ several robustness tests to investigate the sensitivity of our main conclusions. These tests focus on benchmark sensitivity and sub-samples. We run the same tests as in the previous section and below we will touch upon the most important findings.

As mentioned already in section 2.1, indices selection is one of the most important issues in the RBSA. Hence, we pay close attention to the benchmark choice in setting up a Sharpe asset class factor model especially when the correlation coefficient between benchmarks is high. As Berk Sensoy (2009) discussed, over 90% of U.S. equity mutual funds use the S&P or Russell benchmark index. He finds that, over 44% of mutual funds declare the S&P500 as their benchmark and more than 40% use the Russell indices. Kim et al (2005) also use Russell indices as benchmarks for their style analysis. According to Table 1, the correlations between benchmarks are high. For example, the correlation between value and growth benchmarks are 0.92, between small-cap and value is 0.89 and between small-cap and growth benchmarks is 0.87 respectively, this may cause problems. To assess the influence of this choice on our results we replace both the S&P500 value and growth benchmark by the Russell 1000 value and Russell 1000 growth and we consider Russell 2000 as small cap benchmark, which are available from the FactSet Research System Inc.

The use of the Russell indices does not have any impact on our previous results, except for growth/income funds, which we will consider in some more detail hereafter. As table 12 shows, Growth/Income funds now show a higher exposure to the Growth benchmark (0.46) and a lower exposure to the small cap benchmark (0.09).

The results using the Russell indices also reveal the same percentage of significantly misclassified funds as before. This is yet another example where the use of statistical tests on the portfolio weights plays an important role. Point estimates alone overstate the sensitivity for alternative benchmarks. Our overall conclusions on the reduction of significantly misclassified funds and objective gaming remain unchanged.

**Table 12: Results Sharpe asset class factor model**

**Panel A:** Estimated style weights for Alive Portfolio

| Objective | Bond | Cash | Value | Growth | Small cap | $R^2$ |
|---|---|---|---|---|---|---|
| Growth | 0.00 | 0.00 | 0.15*** | 0.63*** | 0.21*** | 0.98 |
| Growth/Income | 0.01*** | 0.10*** | 0.35*** | 0.46*** | 0.09*** | 0.99 |
| Income | 0.01*** | 0.07*** | 0.53*** | 0.35*** | 0.03*** | 0.99 |
| Small cap | 0.00 | 0.00 | 0.00 | 0.19*** | 0.81*** | 0.98 |

**Panel B:** 95% Confidence intervals for style weights for Alive Portfolio

| Objective | Bond | Cash | Value | Growth | Small cap |
|---|---|---|---|---|---|
| Growth | [0.00 – 0.00] | [0.00 – 0.05] | [0.06 – 0.24] | [0.55 – 0.71] | [0.16 – 0.26] |
| Growth/Income | [0.00 – 0.02] | [0.08 – 0.15] | [0.28 – 0.41] | [0.40 – 0.51] | [0.04 – 0.13] |
| Income | [0.00 – 0.01] | [0.05 – 0.11] | [0.48 – 0.58] | [0.30 – 0.41] | [0.00 – 0.03] |
| Small cap | [0.00 – 0.00] | [0.00 – 0.00] | [0.00 – 0.00] | [0.15 – 0.34] | [0.75 – 0.87] |

*Notes* This table presents the parameter estimates of the Sharpe return-based model for six value weighted portfolios of funds. In panel A estimated style weights are given. Each row deals with one particular investment objective, where the elements in columns 3 to 7 report the estimated style weights. Panel B reports the 95% confidence intervals for all estimated style weights. Because of the constraints on the parameters these have been constructed by bootstrapping.

*** Significantly different from zero at the 1 % level

** Significantly different from zero at the 5 % level

* Significantly different from zero at the 10% level

With Significance in Panel A based on confidence intervals reported in Panel B.

As a second robustness check we divide our sample period into two equal periods to investigate the consistency of our previous observations.

Cooper (2005) finds that mutual funds change their names to take advantage of current hot investment styles. Hence, based on the changes of superiority of Value and Growth indices during the time, we divide our sample period into two equal sub-periods to investigate whether fund managers strategically invest in the "Hot" style and away from the current "Cold" style regardless of their stated investment objective to attract more investors.

The first sub-period runs from July 2003 – July 2008 and the second sub-period from August 2008 – December 2014. In the first period, when the Value index in most times beats the Growth index, the annual mean difference return of Growth index and Value is -1.36% and in the second period, when the Growth index shows on average a better performance than the

Value index, the mean difference return is equal to 1.71%, hence the Growth index is becoming hotter than Value. Going from first period to second period, for both value and growth funds the exposure to the value benchmark decreases and the exposure to the growth benchmark increases. Based on point estimates alone the results in an increase of misclassified income funds and a decrease in the misclassified growth funds. The result also shows that Growth/Income funds tend to invest in Growth equities in case of the second period. Therefore, income mutual fund managers tend to strategically invest in the current “hot” style regardless their stated investment objective.

**Table 13: Results Sharpe asset class factor model**

**Panel A:** Estimated style weights for Alive Portfolio- from July 2003 – July 2008

| Objective | Bond | Cash | Value | Growth | Small cap | $R^2$ |
|---|---|---|---|---|---|---|
| Growth | 0.00* | 0.00 | 0.15* | 0.58* | 0.27* | 0.97 |
| Growth/Income | 0.02* | 0.07* | 0.36* | 0.43* | 0.11* | 0.98 |
| Income | 0.04* | 0.05 | 0.56* | 0.30* | 0.04* | 0.97 |
| Small cap | 0.00* | 0.00 | 0.00 | 0.14* | 0.86* | 0.98 |

**Panel B:** Estimated style weights for Alive Portfolio- from August 2008 – December 2014

| Objective | Bond | Cash | Value | Growth | Small cap | $R^2$ |
|---|---|---|---|---|---|---|
| Growth | 0.00* | 0.00 | 0.15* | 0.66* | 0.18* | 0.98 |
| Growth/Income | 0.00* | 0.10* | 0.33* | 0.48* | 0.07* | 0.98 |
| Income | 0.01* | 0.08* | 0.49* | 0.38* | 0.03* | 0.99 |
| Small cap | 0.00 | 0.00 | 0.00 | 0.21* | 0.79* | 0.98 |

Notes this table reports the estimated style weights for Alive Portfolio. The significance levels are abbreviated with asterisks: *, **, and *** denote significance at the 15%, 10%, and 5% levels, respectively.

Finally we attempt to examine the impact of financial crisis on the objective misclassification. Hence, we focus on the specific period of time which is started from April 2008 to February 2009. As table 14 shows, the Growth funds that they were never sensitive to Cash and Bond indices, during the financial crisis, Cash and Bond factor loadings had been statistically significant.It can show that during that time fund managers tend to have a conservative investment strategy by selling their underlying assets and buying bonds and keeping more Cash.

**Table 14: Results Sharpe asset class factor model**

**Panel A:** Estimated style weights for Alive Portfolio- from July 2003 – July 2008

| Objective | Bond | Cash | Value | Growth | Small cap | $R^2$ |
|---|---|---|---|---|---|---|
| Growth | 0.00*** | 0.01*** | 0.05*** | 0.78*** | 0.17*** | 0.98 |
| Growth/Income | 0.01*** | 0.10*** | 0.19*** | 0.60*** | 0.11*** | 0.99 |
| Income | 0.02*** | 0.08*** | 0.41*** | 0.45*** | 0.04*** | 0.99 |
| Small cap | 0.00 | 0.00 | 0.00 | 0.41*** | 0.59*** | 0.99 |

Notes this table reports the estimated style weights for Alive Portfolio. The significance levels are abbreviated with asterisks: *, **, and *** denote significance at the 15%, 10%, and 5% levels, respectively.

## Conclusion

Prior studies show that before 2000, actual investment styles of over 50% of U.S. equity mutual funds did not match their stated investment objective, and more than 30% of the funds were severely misclassified. We observe a drop in misclassification behavior after 2004. Our results for the period July 2003 to December 2014 show that the number of misclassification and severe misclassification decreases dramatically to 25% and 18%, respectively. One possible reason to explain this diminution can be employing inaccurate methods to figure out the number of misclassified funds in the past. Hence, we propose an alternative approach that allows us to infer confidence intervals for the style coefficients.

We next find that on average the annual return of misclassified funds is 0.86% lower than the return on well-classified funds and when we take the significance of estimated style weights into account, this number increases to 1.08%. In addition, based on various fund performance measures, we find that misclassified mutual funds underperform the well-classified funds. Thus on both an unadjusted and risk-adjusted basis, misclassified funds underperform their well-classified peers in each investment style categories.

We also find the common characteristics of misclassified mutual funds that help us to identify the motivation behind misclassification behavior. We show that on average misclassified funds are typically belong to small-sized funds, weak prior performance, retail funds, high-expenses funds and high trading cost funds. In conclusion, misclassified funds are typically distressed funds that appear to deviate from their benchmarks to reverse current performance and we argue that the deviation from stated investment style is a kind of zero sum game.

Our result also shows that fund managers of misclassified funds are prone to hug the current hot indices regardless of their stated investment objectives, it suggests that misclassification cannot be based on simple negligence or unintentional behaviour.

# Appendix

## Notation

Recall the model that was set-up in section 2 is of the form

$$r_{it} = \alpha_i + \sum_{j=1}^{N} \beta_{ij} I_{jt} + \varepsilon_{it} \qquad t = 1, \dots, T \tag{1A}$$

$$\sum_{j=1}^{N} \beta_{ij} = 1 \tag{2A}$$

$$\beta_{ij} \geq 0 \qquad j = 1, \dots, N \tag{3A}$$

In matrix notation this model is given by

$$Y = X\beta + u \tag{4A}$$

$$j'\beta = 1 \tag{5A}$$

$$\beta_k \geq 0 \qquad k = 2, \dots, K \tag{6A}$$

where

$$Y = \begin{pmatrix} r_{i1} \\ \vdots \\ r_{iT} \end{pmatrix} \tag{7A}$$

$$X = \begin{pmatrix} 1 & I_{11} & \cdots & I_{N1} \\ \vdots & \vdots & & \vdots \\ 1 & I_{1T} & \cdots & I_{NT} \end{pmatrix} \tag{8A}$$

$$\beta = \begin{pmatrix} \alpha_i \\ \beta_{i1} \\ \vdots \\ \beta_{iN} \end{pmatrix} \tag{9A}$$

and

$$j' = \begin{pmatrix} 0 & 1 & \cdots & 1 \end{pmatrix} \tag{10A}$$

The dimensions are for Y $(T \times 1)$, for X $(T \times K)$, for b $(K \times 1)$, for $j$ $(K \times 1)$ and for u $(T \times 1)$. For notational convenience the subscript i has been suppressed. In the following we first show the estimation results for the unconstrained model, then for the Lagrange model in which we also take account for the equality constraint, and finally the model with both equality and inequality constraints is tackled. We show that the Lagrange estimator $(b_L)$ can be written in terms of the unconstrained estimator $(b_U)$, and that the Kuhn-Tucker estimator $(b_{KT})$ can be written in terms of the Lagrange estimator (and therefore also in terms of the unconstrained estimator).

**Unconstrained model**

The unconstrained estimator minimizes the sum of squares in equation (4A) and is given by

$$b_U = (X'X)^{-1} X'Y \tag{11A}$$

and the asymptotic distribution of the parameter estimates is given by

$$b_U \sim N(\beta, V_{bU}) \tag{12A}$$

**Lagrange model**

The Lagrange estimator minimizes the sum of squares in equation (4A) subject to the equality constraint in equation (5A) and is given by

$$b_L = (I_K - Pj')b_U + P \tag{13A}$$

where $I_K$ is the $(K \times K)$ identity matrix and the matrix P is given by

$$P = (X'X)^{-1} j\left[j'(X'X)^{-1} j\right]^{-1} \tag{14A}$$

<u>**Proof:**</u>

The Lagrangian is given by

$$L = (Y - X\beta)'(Y - X\beta) - \lambda(j'\beta - 1) \tag{15A}$$

where $\lambda$ denotes the Lagrange multiplier. The first order condition reads

$$\frac{\partial L}{\partial \beta} = 0: \quad -2X'(Y - X\beta) - j\lambda = 0 \tag{16A}$$

So, it follows that

$$b_L = (X'X)^{-1} X'Y + \tfrac{1}{2}(X'X)^{-1} j\lambda = b_U + \tfrac{1}{2}(X'X)^{-1} j\lambda \tag{17A}$$

From (5A) it follows that

$$1 = \imath b_L = \imath b_U + \tfrac{1}{2} j'(X'X)^{-1} j\lambda \tag{18A}$$

which implies that

$$\lambda = 2\left[j'(X'X)^{-1} j\right]^{-1}(1 - j'b_U) \tag{19A}$$

Substitution of (19A) in (17A) leads to the estimator in (13A).

**Kuhn-Tucker model**

The Kuhn-Tucker model minimizes the sum of squares in equation (4A) subject both to the equality constraint in (5A) and the inequality constraints in (6A). The most straightforward solution to a Kuhn-Tucker problem is to consider it as $2^K$ Lagrange sub-problems, in which the sum of squares is minimized subject to each possible combination for which the inequality constraint is either binding or non-binding. The Lagrange sub-problem that leads to the parameter estimates with the lowest sum of squares and that also meets all restrictions also leads to the Kuhn-Tucker estimator.

The Kuhn-Tucker estimator $(b_{KT})$ is given by the estimator of all sub-problems $(b_S)$ that minimizes the sum of squares and fulfils all restrictions. Let S be the matrix that represents the binding inequality constraints, i.e. the associated equality constraint reads

$$S\beta = 0 \tag{20A}$$

For example, the case where the second and the third parameter are binding is represented by the $(2 \times K)$ matrix

$$S = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \end{pmatrix} \tag{21A}$$

The solution of the sub-problem of minimizing the sum of squares in equation (4A) subject to both the equality constraint in (6A) and in (20A) is given by

$$b_S = \left\{I_K - VS'(SVS')^{-1}S\right\}P + \left\{I_K - VS'(SVS')^{-1}S\right\}[I_K - P]b_U \tag{22A}$$

where

$$V = \left(I_K - (X'X)^{-1} j\left\{j'(X'X)^{-1} j\right\}j'\right)(X'X)^{-1} = (I_K - Pj')(X'X)^{-1} \tag{23A}$$

and the associated Kuhn-Tucker estimator reads

$$b_{KT} = \min_S \left\{(Y - Xb_S)'(Y - Xb_S) \mid j'b_S = 1; Sb_S = 0\right\} \tag{24A}$$

**<u>Proof:</u>**

The Lagrangian is given by

$$L = (Y - X\beta)'(Y - X\beta) - \lambda(j'\beta - 1) - \mu(S\beta) \tag{25A}$$

where $\lambda$ and $\mu$ denote the Lagrange multipliers. The first order condition reads

$$\frac{\partial L}{\partial \beta} = 0: \quad -2X'(Y - X\beta) - j\lambda - S'\mu = 0 \tag{26A}$$

So, it follows that

$$b_S = b_U + \tfrac{1}{2}(X'X)^{-1} j\lambda + \tfrac{1}{2}(X'X)^{-1} S'\mu \tag{27A}$$

From equation (5A) it follows that

$$1 = j'b_S = j'b_U + \tfrac{1}{2}\left\{j'(X'X)^{-1} j\right\}\lambda + \tfrac{1}{2} j'(X'X)^{-1} S'\mu \tag{28A}$$

Solving for $\lambda$ gives

$$\lambda = 2\left[j'(X'X)^{-1} j\right]^{-1}(1 - j'b_U) - \left[j'(X'X)^{-1} j\right]^{-1} j'(X'X)^{-1} S'\mu \tag{29A}$$

Substitution of (29A) in (27A) gives

$$b_S = b_L + \tfrac{1}{2} VS'\mu \tag{30A}$$

Now, use the relation in equation (20A), to arrive at

$$0 = Sb_S = Sb_L + \tfrac{1}{2}(SVS')\mu \tag{31A}$$

Solving for $\mu$ gives

$$\mu = -2(SVS')^{-1} Sb_L \tag{32A}$$

Substitution of (32A) in (30A) gives

$$b_S = \{I - VS'(SVS')^{-1} S\} b_L \tag{33A}$$

Substitution of (13A) in (33A) gives the required result.

## References


Anderson E., L.P. Hansen, and T.J. Sargent, 2000, Robustness, detection and the price of risk, Working Paper University of North Carolina, University of Chicago and Stanford University

Andrews, D., 1997, Estimation when a Parameter is on a Boundary: Theory and Applications, Discussion Paper Cowles Foundation, Yale

Andrews, D., 2000, Inconsistency of the Bootstrap when a Parameter is on a Boundary, Econometrica 67, 1341 - 1383

Barberis N., 2000, Investing for the long run when returns are predictable, Journal of Finance 55, 225-264

Brown S., and W. Goetzmann, 1997, Mutual Fund Styles, Journal of Financial Economics 43, 373-399

DiBartolomeo D., and E. Witkowski, 1997, Mutual Fund Misclassification: Evidence based on Style Analysis, Financial Analysts Journal Sept/Oct, 32-43

Gouriéroux, Ch., Holly A., and A. Montfort, 1982, Likelihood Ratio Test, Wald Test and Kuhn-Tucker Test in Linear Models with Inequality Constraints on the Regression Parameters, Econometrica 50, 63-80

Gouriéroux, Ch., and A. Montfort, 1995, Statistics and Econometric Models, volume 2, Cambridge University Press, 536p

Ippolito R., Consumer Reaction to Measures of Poor Quality: Evidence from the Mutual Fund Industry, Journal of Law and Economics, 1992, 45-70

Judge, G.G. and T. Takayama, 1966, Inequality Restrictions in Regression Analysis, Journal of American Statistical Association 61, 166-181

Kodde D.A., and F.C. Palm, 1986, Wald Criteria For Jointly Testing Equality and Inequality Restrictions, Econometrica 54, 1243-1248

Kim M., R. Shukla and M. Tomas, 2000, Mutual Fund Objective Misclassification, Journal of Economics and Business 52, 309-323

Kim T-H, D. Stone and H. White, 2000, Asymptotic and Bayesian Confidence Intervals for Sharpe Style Weights, Discussion Paper No. 00/18 University of Nottingham

Lobosco A., and D. DiBartolomeo, 1997, Approximating the Confidence Intervals for Sharpe Style Weights, Financial Analysts Journal July/August, 80-85

Maenhout, P. J. 2004, Robust portfolio rules and asset pricing. Review of financial studies, 17(4), 951-983.

Sharpe W., 1988, Determining a Fund’s Effective Asset Mix, Investment Management Review vol 2/6 December, 59-69

_________, 1992, Asset Allocation: Management Style and Performance Measurement, Journal of Portfolio Management 46, 7-19

Meier, I., & Schaumburg, E. 2004, Do funds window dress? Evidence for US domestic equity mutual funds. Unpublished working paper, HEC Montreal and Kellogg School of Management.

Kim, T. H., White, H., & Stone, D. (2005). Asymptotic and Bayesian confidence intervals for Sharpe-style weights. Journal of Financial Econometrics, 3(3), 315-343.

Petajisto, A. (2013). Active share and mutual fund performance. Financial Analysts Journal, 69(4), 73-93.

Agarwal, V., Gay, G. D., & Ling, L. (2014). Window dressing in mutual funds.Review of Financial Studies, hhu045.

Mason, A., McGroarty, F., & Thomas, S. (2012). Style analysis for diversified US equity funds. Journal of Asset Management, 13(3), 170-185.

ter Horst, J. R., Nijman, T. E., & de Roon, F. A. (2004). Evaluating style analysis. Journal of Empirical Finance, 11(1), 29-53.

Rekenthaler, J., Gambera, M., & Charlson, J. (2006). Estimating portfolio style in US equity funds: A comparative study of portfolio-based fundamental style analysis and returns-based style analysis. The Journal of Investing, 15(3), 25-33.

Almazan, A., Brown, K. C., Carlson, M., & Chapman, D. A. (2004). Why constrain your mutual fund manager?. Journal of Financial Economics, 73(2), 289-321.

Geweke, J. (1986). Exact inference in the inequality constrained normal linear regression model. Journal of Applied Econometrics, 1(2), 127-141.

Huang, J., Sialm, C., & Zhang, H. (2011). Risk shifting and mutual fund performance. Review of Financial Studies, hhr001.

Huang, J., Wei, K. D., & Yan, H. (2007). Participation costs and the sensitivity of fund flows to past performance. The Journal of Finance, 62(3), 1273-1311.

Sirri, E. R., & Tufano, P. (1998). Costly search and mutual fund flows. The journal of finance, 53(5), 1589-1622.

Brown, S. J., & Goetzmann, W. N. (2001). Hedge funds with style (No. w8173). National Bureau of Economic Research.

Nanda, V., Z. J. Wang, and L. Zheng. 2004. Family Values and the Star Phenomenon: Strategies of Mutual Fund Families. Review of Financial Studies 17:667–98.

Li, W., and A. Tiwari. 2006. On the Consequences of Mutual Fund Tournaments. University of Iowa.

Hu, P., J. R. Kale, M. Pagani, and A. Subramanian. 2010. Fund Flows, Performance, Managerial Career Concerns, and Risk-taking. Georgia State University and San Jose State University.

Chalmers, J. M., R. M. Edelen, and G. B. Kadlec. 1999. An Analysis of Mutual Fund Trading Costs. University of Oregon, University of Pennsylvania, and Virginia Tech.

Huang, J., K. D. Wei, and H. Yan. 2007. Participation Costs and the Sensitivity of Fund Flows to Past Performance. Journal of Finance 62:1273–311

Gil-Bazo, J., and P. Ruiz-Verdu. 2009. The Relation Between Price and Performance in the Mutual Fund Industry. Journal of Finance 64:2153–83.

Sensoy, B. A. (2009). Performance evaluation and self-designated benchmark indexes in the mutual fund industry. Journal of Financial Economics, 92(1), 25-39.

Cooper, M. J., Gulen, H., & Rau, P. R. (2005). Changing names with style: Mutual fund name changes and their effects on fund flows. The Journal of Finance, 60(6), 2825-2858.

Berk, J., & Van Binsbergen, J. H. (2014). Measuring skill in the mutual fund industry. Available at SSRN 2038108.

Nanda, V. K., Wang, Z. J., & Zheng, L. (2009). The ABCs of mutual funds: On the introduction of multiple share classes. Journal of Financial Intermediation,18(3), 329-361.

Kacperczyk, M., Sialm, C., & Zheng, L. (2008). Unobserved actions of mutual funds. Review of Financial Studies, 21(6), 2379-2416.

Ferruz Agudo*, L., & Vicente Gimeno, L. A. (2005). Effects of multicollinearity on the definition of mutual funds' strategic style: the Spanish case. Applied Economics Letters, 12(9), 553-556.

Andrews, D. W., & Buchinsky, M. (2000). A Three- step Method for Choosing the Number of Bootstrap Repetitions. Econometrica, 68(1), 23-51.

Cameron, A. C., & Trivedi, P. K. (2005). Microeconometrics: methods and applications. Cambridge university press.

Davidson, R., & MacKinnon, J. G. (2000). Bootstrap tests: How many bootstraps?. Econometric Reviews, 19(1), 55-68.

Swinkels, L., & Van Der Sluis, P. J. (2006). Return-based style analysis with time-varying exposures. The European Journal of Finance, 12(6-7), 529-552.

Berk, J. B., & Walden, J. (2013). Limited capital market participation and human capital risk. Review of Asset Pricing Studies, 3(1), 1-37.

Chen, J., Hong, H., Huang, M., & Kubik, J. D. (2004). Does fund size erode mutual fund performance? The role of liquidity and organization. American Economic Review, 94(5), 1276-1302.

Elton, E. J., Gruber, M. J., & Blake, C. R. (2001). A first look at the accuracy of the CRSP mutual fund database and a comparison of the CRSP and Morningstar mutual fund databases. The Journal of Finance, 56(6), 2415-2430.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. Journal of financial economics, 33(1), 3-56.

Carhart, M. M. (1997). On persistence in mutual fund performance. The Journal of finance, 52(1), 57-82.

Pástor, L., & Stambaugh, R. F. (2003). qLiquidity Risk and Expected Stock Retums. rThe Journal of Political Economy, 111, 642-685.