When Mutual Fund Names Misinform*

Anne-Florence Allard,† Jonathan Krakow,‡ and Kristien Smedts§

*We are grateful for comments from Christoph Basten, Marc Chesney, Hans Degryse, Mike Gallmeyer, Thorsten Hens, Leonardo Iania, Steven Ongena, Randy Priem, Silvina Rubio, Timo Schäfer, Paul Smeets, Gertjan Verdickt, Alexander Wagner, Gunther Wuyts, and Alexandre Ziegler, and seminar participants at the University of Bristol, and the University of Zurich. Anne-Florence Allard acknowledges financial support from the Research Foundation - Flanders (FWO).

†School of Accounting and Finance, University of Bristol, UK & Department of Accounting, Finance & Insurance, KU Leuven, Belgium. Email: anneflorence.allard@bristol.ac.uk

‡Department of Banking and Finance, University of Zurich, Switzerland. Email: jonathan.krakow@bf.uzh.ch

§Department of Accounting, Finance & Insurance and LRisk, KU Leuven, Belgium. Email: kristien.smedts@kuleuven.be

When Mutual Fund Names Misinform

Abstract

Mutual funds often use their name to inform about the investment style that is pursued. We document that a significant proportion of mutual fund names provides inaccurate information, ie the fund's investment style does not align with the given name. Funds that deviate from the investment style suggested in their name performed worse, had lower fund inflows and took more risk than their competitors. Evidence shows that it is, in particular, the tournament character of the fund industry that causes fund name deviations. In addition, the risk-return trade-off deteriorates following a misnaming practice. Moreover, we document that investors experience difficulties in responding in a timely manner to this misleading name information, especially when the fund name includes the term "growth" or "value".

JEL classification: G11, G14, G18

Keywords: Mutual funds, fund names, cluster analysis, moral hazard, fund performance

1. Introduction

Do mutual fund names convey reliable information? In this paper, we shed light on this question by studying mutual fund names' alignment with mutual funds' *actual* investment styles.

At year-end 2018, there were more than 119,000 regulated funds worldwide having more than US $46 trillion in total net assets (Investment Company Institute, 2019). Investors thus face a broad menu of investment options, and to simplify the selection process, they often rely on information or heuristics summarising key fund characteristics. For instance, investors (over-)weight past performance, explicit fees or base their decision on qualitative information such as the fund's name or its ticker symbol (see, eg Sirri and Tufano, 1998; Jain and Wu, 2000; Barber, Odean, and Zheng, 2005; Cooper, Gulen, and Rau, 2005; Espenlaub, ul Haq, and Khurshed, 2017).[1]

As a result, fund managers often include salient information in the fund's name – such as the asset class, eg equity/bonds, or the investment style of the fund, eg value/growth, or small/mid/large-cap. Providing such information in the name should facilitate the selection process of investors, for sure when they are less sophisticated. However, what if fund managers do not truthfully inform—or even intentionally misinform—about the strategy in the fund's name?

In this paper, we investigate this phenomenon of inaccurate fund names. Using cluster analysis coupled with a return-based analysis, we are able to show the prevalence of inaccurate fund names, ie a misalignment between the fund's investment style and name. Subsequently, we aim to shed light on the reasons and consequences of inaccurate fund names. This analysis is motivated by the tournament hypothesis of Brown, Harlow, and Starks (1996)

[1] Thaler (2016) provides anecdotal evidence of investors' behavior regarding a closed-end mutual fund. A fund having the ticker symbol "CUBA" and investing in the Caribbean traded historically (like most closed-end funds) at a 10 to 15% discount relative to the net assets value. However, in December 2014, when President Obama announced to relax conditions for US firms to do business in Cuba, this fund experienced a significant increase in the price of more than 70% without any change in the net asset value. Thus, investors overweighted the qualitative information based on the ticker symbol and interpreted it as a relation to the country Cuba.

who consider the mutual fund market as a competitive tournament in which funds with comparable investment objectives compete against each other. They show that relative winners within a fund category experience higher fund inflows relative to their competitors. Moreover, higher inflows result in higher compensation for fund managers (Chevalier and Ellison, 1997). Therefore, in an attempt to become relative winners, mutual fund managers might decide to follow a deviating strategy. This mainly plays a role at year end, when the fund's performance is recorded and the compensation is determined. Based on this relationship, we expect that relative losers over the course of a year will increase their efforts to switch to the relative winners' side by the end of the year. One strategy to do so is to deviate from the investment style suggested by the fund's name and change the portfolio's exposure profile. This effect is amplified due to an asymmetric return-flow relationship, where relative winners gain inflows, but relative losers do not experience outflows (see Brown et al., 1996; Chevalier and Ellison, 1997). Due to this asymmetric pattern of the flow-performance relationship and due to the tournament setting, managers of funds with relatively lower returns over a year thus have incentives to change the investment style of the fund by the end of the year. The fund name's inaccuracy therefore emerges from a principal-agent conflict, in which the fund manager aims to maximize their compensation.

Finally, we analyze the consequences of improper naming practices. We study the fund's subsequent return and risk to shed light on the success of the strategy to deviate. We also analyze the investors' subsequent reaction by investigating fund in- and outflows.

Over a period spanning from 2010 up to 2018, we document that a significant fraction of US equity mutual funds provides inaccurate naming information: 33% of US equity mutual funds have, at least once in their life-cycle, an inaccurate name. This inaccuracy especially appears during the last quarter of a calendar year. However, misnaming often occurs for multiple periods over the life-cycle.

To explore this finding further, we study which fund characteristics can be linked to this misnaming practice. We show that mutual funds, prior to their misnaming, underperform in many dimensions compared to accurately named funds. In particular, the degree of inaccuracy is higher when the fund received lower inflows and was exposed to more idiosyncratic risk. Our results also highlight the tournament hypothesis and confirm the potential principal-agent conflict (Brown et al., 1996). When funds are relative losers during the first three quarters of a year, they tend to deviate more from the style suggested by their name in the last quarter of the calendar year, resulting in inaccurate fund naming.

Importantly, we document that the above strategy is not successful: focusing on relative performance, funds that deviate from their name-suggested investment style do not end-up being relative winners. Thus, a deviation to improve one's performance rank does not seem to be effective, on average. Moreover, while the risk-return trade-off deteriorates following a misnaming practice, investors do not significantly reduce their investments in such funds in a timely manner, especially when the fund name includes the term *growth* or *value*. This finding is in line with the asymmetric relationship of returns and flows (see, eg Sirri and Tufano, 1998) and highlights the difficulties that investors experience in noticing such inaccurate information. Based on these findings, we argue for the need for a stricter name regulation.

The importance of mutual fund names in investors' investment decisions has long been recognized by the Security and Exchange Commission (SEC). For instance, in 2001, the SEC stated that "the name of an investment company may communicate a great deal to an investor" (SEC, 2001). This recognition resulted in Rule 35d-1, introduced in July 2001, to regulate mutual fund names. Hence, to prevent misleading information in the name, the rule requires funds that mention an asset class in their name to invest at least 80% of the portfolio in that asset class. Following the introduction of this naming rule, a significant fraction of funds had to make name adjustments, as reported by Cooper et al. (2005).

Rule 35d-1 is transparent and strict concerning the asset classes, the sector, or the region in which a fund invests, whether the distribution is exempt from the income tax, and whether the fund shares are guaranteed or approved by the United States government. However, the rule is not strictly enforceable when the fund name indicates a particular focus on asset size, for example, on *small* or *large* capitalized companies. Instead, the authorities require that a fund uses "reasonable definitions of such terms", leaving room for interpretation. Some other aspects of the name are even left completely unregulated, such as a name related to a typical investment strategy (*growth* or *value*).[2]

Aware of these possible loopholes in the regulation, the SEC issued a press release in March 2020 requesting public comments about the effectiveness of Rule 35d-1. In this press release, authorities highlight the concerns linked to fund names referring to the size and investment strategy dimensions (SEC, 2020). Strongly motivated by this concern, our analysis therefore focuses on fund names referring to the size dimension (*small* or *large*) and the investment strategy dimension (*value* or *growth*).

In this paper, we do not assume that a mutual fund name is exclusive in the way that it indicates investments *only* in the suggested asset class or *only* in stocks corresponding to the specified investment style, ie a fund called *small* might also invest to a certain (much lesser) degree in *large* capitalized companies. Mutual fund names are only one of the information sources that investors can consult, and it should not be the only one. However, as the SEC states: "fund names are often the first piece of information investors see, and they can have a significant impact on an investment decision" (SEC, 2020). Hence, the name of a mutual fund introduces essential information to potential investors, and therefore, should not be misleading.

Regarding the existing literature, our paper builds upon the literature on inaccurate information in the mutual fund industry in general. In particular, there is a strand of literature that investigates mismatches between what mutual funds claim to do and what they actually do: mismatches between investment styles in funds' objective statements and actual investment styles (Bams, Otten, and Ramezanifar, 2017; Brown and Goetzmann, 1997; Kim, White, and Stone, 2005; Mason, Mcgroarty, and Thomas, 2012), between stated

[2]All names are still subject to the prohibition on misleading names (SEC, 2001).

investment objectives in general and actual objectives (Kim, Shukla, and Tomas, 2000), and between the stated benchmark index and actual investment style (DiBartolomeo and Witkowski, 1997; Mateus, Mateus, and Todorovic, 2019; Sensoy, 2009) or actual holdings (Cremers and Petajisto, 2009). We contribute to this literature by providing evidence on mismatches related to the fund's name, the very first qualitative information that investors see when they invest in a mutual fund, and its actual investment style.

Second, we also contribute to the strand of literature that investigates the role of mutual fund naming in more detail. For example, Kumar, Niessen-Ruenzi, and Spalt (2015) provide evidence that US investors' decisions are not free of name-induced stereotypes: funds with managers having foreign-sounding names have significantly lower annual fund flows. More closely related to our paper are two studies that report empirical evidence of misinformation related to *changes* in mutual fund names (Cooper et al., 2005; Espenlaub et al., 2017). Cooper et al. (2005) analyze fund name changes from April 1994 to July 2001 and find that funds use name changes strategically. While name-changing funds are not able to improve return-performance, they do experience significant long-term positive abnormal fund inflows. This relationship even holds for a cosmetic name change, ie when the portfolio holdings of the fund are not in line with its new name. Espenlaub et al. (2017) build further upon Cooper et al. (2005) and study fund name changes between 2002 and 2011, after the implementation of the SEC Names Rule in 2001. They distinguish name changes along several dimensions, among which misleadingness which corresponds to the definition of cosmetic name changes in Cooper et al. (2005).

To summarize, our contribution is fourfold. First, we analyze the problem of inaccurate fund naming by taking a broader, more continuous perspective, which is, in comparison to the above literature (Cooper et al., 2005; Espenlaub et al., 2017), not restricted to name change events. We find that the economic problem of inaccurate names is indeed much broader than just triggered by name change events. Second, we provide new insights for regulators by focusing on name dimensions, both strictly, but also less strictly, regulated by the SEC Names Rule SEC (2001). Interestingly, we find fewer mismatches in the size dimension (SMB), which is more strictly regulated by the SEC. In contrast, we find significant mismatches in the strategy dimension (HML), which is less strictly regulated. Third, we show that the misnaming practice can be explained by the fund industry's tournament character. Finally, we analyze the consequences of name misinformation and highlight the importance of name regulation, given investors' limited reaction to inaccurate naming.

The remainder of the paper is organized as follows. Section 2 describes the data collection process and provides descriptive summary statistics. In Section 3, we develop our accuracy classification methodology and derive the corresponding classifications for our sample. The second part of the paper elaborates on the reasons and potential consequences of inaccurate mutual fund names. Section 4 analyzes the funds' motivation to deviate from the investment style suggested by their name and the role played by the mutual fund tournament. Section 5 investigates the consequences of inaccurate fund names. Finally, Section 6 concludes.

2. Data

To analyze potential misleading information in mutual fund names, we construct a history of mutual fund names. To do so, we consult over 400,000 fund prospectuses available on the EDGAR database of the SEC and link the fund names to key financial data obtained from Morningstar.

2.1. EDGAR: Mutual fund names

The starting point to construct a history of mutual fund names is the open-access database EDGAR of the SEC. According to the Securities Acts of 1933 and 1940, investment companies registered with the SEC are required to disclose standardised information on their investment products, eg mutual funds.[3] According to the Securities Acts, funds are

[3]There is a relatively young research stream on information from mutual fund prospectuses (see Abis, 2017; Baghai, Becker, and Pitschner, 2019; Hillert, Niessen-Ruenzi, and Ruenzi, 2016; Kostovetsky and

required to (1) update at least once a year, material fund information and (2) disclose a separate filing whenever there is a significant change in this information. For completeness, we extract name information from *all* filings that investment companies of funds regularly have to provide (see Table A in the Appendix for a detailed list of filing types).

Therefore, beginning in 2010, we extract the names of a fund from each fund prospectus in the EDGAR database.[4] We analyze a total of 418,938 prospectuses that were filed between January 2010 and December 2018. We then aggregate the information from these different filings and construct a monthly fund name history database of 20,973 funds (see Table A in the Appendix for the details).

2.2. Morningstar: Financial data of mutual funds

The names history database is merged with the survivorship bias free mutual fund database of Morningstar that includes information on fund returns, various risk measures, as well as several mutual fund characteristics, such as total assets under management, costs, and security holdings.[5]

In order to analyze investors' reactions, we compute the fund quarterly net flows. Following the methodology of Huang, Sialm, and Zhang (2011), we calculate—based on the fund's total net assets and quarterly gross returns—the quarterly fund flows as:

$$Flow_{i,t} = \frac{TNA_{i,t} - TNA_{i,t-1}(1 + R_{i,t})}{TNA_{i,t-1}(1 + R_{i,t})}, \tag{1}$$

Warner, 2020; Krakow and Schäfer, 2020). Abis (2017) applies machine learning techniques to the strategy section of mutual fund prospectuses to classify funds as either quantitative or discretionary. Baghai et al. (2019) retrieve information on the use of credit ratings from mutual fund prospectuses. Hillert et al. (2016) analyze the tone in mutual fund letters, which they also extract from the same database. Krakow and Schäfer (2020) provide evidence that qualitative information in fund prospectuses is of value for investors.

[4]We choose 2010 since from that year onward, prospectus data is provided in the eXtensible Business Reporting Language (XBRL) format. Therefore, the extraction of relevant sections in a prospectus can be identified due to the XBRLkey structure.

[5]As many funds are offered in multiple share classes, which all belong to the same fund and therefore also have the same fund name, we aggregate the Morningstar share class-level information into fund-level information. This aggregation is done by taking the market value-weighted average (see, eg Kacperczyk, Sialm, and Zheng (2008); Gallaher, Kaniel, and Starks (2015); Choi, Kahraman, and Mukherjee (2016)). The oldest share class is used as the reference for all the other variables for which no aggregation is needed.

where $TNA_{i,t}$ is fund i's total net assets in quarter t and $R_{i,t}$ is the fund's return over the prior quarter.[6] To account for outliers, we winsorize at the 1% level at the bottom and the top parts of the distribution.

While data is mostly available on a monthly basis, we aggregate it to a quarterly frequency. We do this to avoid too much variability in factor loadings when using short samples. Name changes within a quarter are handled in the following way. If a fund changes its name in the first half of the quarter, we assign the new name to the corresponding quarter. If the fund changes its name in the second half, we assign the new name only to the next quarter. Finally, we restrict the sample to equity funds only and exclude funds with less than 1 million assets under management and funds with only one quarter of observations.

2.3. Descriptive statistics

Our final sample consists of 2,126 US equity open-end mutual funds that are linked to 2,669 different names in the period 2010 to 2018. Of those funds, 1,339 funds have a name which either refers to the SMB dimension (having the terms *small* or *large* in the name) or to the HML dimension (having the terms *growth* or *value* in the name), or jointly to both dimensions. Table 1 reveals that the number of funds that refer to a specific style is relatively constant over time. A large fraction of funds refers to the HML dimension only, while most funds that refer to the SMB dimension, also refer to the HML dimension.

Table 2 reports sample summary statistics on various portfolio characteristics. The average fund has assets under management of USD 2,709 million, is 16 years old, and charges a quarterly gross expense ratio of 0.12%. The average quarterly fund flow is 2.19%, the average investor's return (net of costs) over a quarter is 1.18%, and the idiosyncratic risk of a fund in a quarter is 0.34%.

[6] This measure ensures that fund flows cannot be below -100%. Most studies define fund's net flows as $Flow_{i,t} = (TNA_{i,t} - TNA_{i,t-1}(1 + R_{i,t}))/(TNA_{i,t-1})$. However, our approach, following the methodology of Huang et al. (2011), guarantees that fund outflows will not be below -100%. The Pearson correlation coefficient between the two measures is 0.998.

Table 1: **Description of mutual fund names**

Table 1 shows the number of fund names in each year that indicate a fund's investment style. Funds are defined as having a style name if one of the following style identifiers appears in their name: "large/lg/blue chip", "small/sml/sm", "growth/grth/gr", and "value/val". The sample consists of all US open-end equity mutual fund names which are retrieved from the SEC database and linked to the Morningstar data over the 2010 to 2018 period.

Dimension	2010	2011	2012	2013	2014	2015	2016	2017	2018
Small	122	131	111	114	116	129	131	144	147
Small & Growth	70	75	60	59	61	63	63	67	70
Small & Value	84	93	76	78	80	85	88	86	86
Large	62	58	46	38	38	42	48	53	60
Large & Growth	79	83	63	56	55	49	48	50	52
Large & Value	77	78	63	57	52	51	54	55	56
Growth	273	281	212	209	204	206	202	217	212
Value	203	215	174	160	151	155	155	167	165
Other	591	612	519	492	472	473	495	553	560
Total	1561	1626	1324	1263	1229	1253	1284	1392	1408

3. Inaccurate mutual fund names

To answer our research questions, we first analyze the prevalence of inaccurate naming practices by studying the alignment between mutual fund names and investment styles. To do so, we rely on a return-based analysis (see, eg Cooper et al., 2005).[7] The motivation for such return-based method is threefold. First, funds' returns provide timely information about the funds' investment style as it is available at a daily frequency, while portfolio holdings data is only updated quarterly. Second, and related, this information is much more sensitive to changes in investment style than the information about portfolio holdings (see, eg ter Horst, Nijman, and de Roon, 2004). Third, while mutual funds report their holdings themselves, returns are publicly observable and are therefore exempt from possible reporting bias.

[7] Cooper et al. (2005) use such a method to identify inaccurate name changes, ie a fund name which refers after a name change to an investment style that does not correspond to the portfolio holdings.

Table 2: **Characteristics of the funds in the sample**

Table 2 shows summary statistics (mean, standard deviation, median, first (p1) and last (p99) percentile. All fund characteristics are defined in detail in Table D. The variables *Fund Flow* and *Expense Ratio* are winsorized at the bottom and the top percentile. The variable *Age* is calculated based on the IPO of the oldest share class. The variable *Idiosyncratic Risk* is the standard deviation of the residual of the Carhart (1997) four-factor model.

Variable	Mean	Std	p1	Median	p99
Fund Flow (in %)	2.19	14.06	-29.61	0.52	81.17
Fund Size (in mn. $)	2709.93	14,568.09	1.53	372.75	43,069.68
log(Fund Size (in mn. $))	19.60	2.18	14.24	19.74	24.49
log(Company Size)	21.80	2.11	16.06	22.07	25.87
Age (in Years)	16.21	12.64	0.83	14.25	74.38
log(Age (in Years))	2.49	0.86	-0.18	2.66	4.31
Return (in %)	1.18	5.06	-18.26	1.15	13.28
Risk (in %)	0.34	0.48	0.04	0.21	2.35
Expense Ratio (in %)	0.12	0.09	0.01	0.09	0.50
Turnover Ratio (in %)	68.77	95.33	2.00	46.61	500.35
Cash Proportion (in %)	2.86	8.64	-0.98	1.86	28.18
FI Proportion (in %)	0.55	7.93	0.00	0.00	7.84
Equity Proportion (in %)	95.95	8.84	58.02	97.71	100.26
Holdings	216.80	404.57	17.33	90.00	2274.00

For each fund, for each quarter, we estimate the four-factor model of Carhart (1997) on daily returns. The factors are obtained from Kenneth French's website.[8] Specifically, we estimate the following model for each quarter:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_{R,i}(R_{M,t} - R_{f,t}) + \beta_{S,i}SMB_t + \beta_{H,i}HML_t + \beta_{M,i}MOM_t + \epsilon_{i,t}, \qquad (2)$$

where $R_{i,t}$ is the return of fund i on day t, and $R_{f,t}$ the risk-free rate. $R_{M,t}$, SMB_t, HML_t and MOM_t denote the market, size, value-growth, and momentum returns on day t, respectively.

[8] We thank Kenneth French for providing data on these factors. For more details, see https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

Thus, as a result, we obtain quarterly factor loadings for each fund. Next, to determine inaccurate fund names, we focus on two of these loadings: $\beta_{S,i}$ and $\beta_{H,i}$. In particular, we use $\beta_{S,i}$ to identify a name referring to the size (SMB) dimension, ie using the terms *small* or *large*. Similarly, we use $\beta_{H,i}$ to identify names referring to the investment strategy (HML) dimension, ie including the terms *growth* or *value*.

For each of these two factor loadings, we use cluster analysis to sort funds into two investment style clusters according to their similarity in $\beta_{S,i}$ or $\beta_{H,i}$. The goal of this cluster analysis is to obtain a high similarity in β between funds in the same group and a low similarity between funds in different groups (see, eg Jain, 2010). Based on the classification of a fund in a group, we can then determine whether the name information is in line with the investment style cluster, and thus whether the name is accurate or not.

Cluster analysis has already been used in the context of mutual funds. For example, Brown and Goetzmann (1997) and Mason et al. (2012) use the most popular algorithm, K-means (Jain, 2010), to classify funds into different style groups. The goal of the K-means algorithm is to allocate funds to a pre-specified number of groups by minimizing the squared error between the group's empirical mean and the funds in this group (Jain, 2010). Despite being intuitive and widely used, this clustering technique has one undesirable property: it is a *hard* assignment, meaning that each fund is assigned to a single cluster. An extension of the K-means algorithm, called fuzzy C-means (Dunn, 1973; Bezdek, 1981), overcomes this by allowing each fund to be part of every group, but with different degrees of membership. For example, in the case of two groups, the fuzzy C-means algorithm identifies the degree of membership of a fund to the first group (say, 0.2) and the second group (say, 0.8). Since the degrees of membership sum up to one, we can interpret them as the likelihood that a fund belongs to each of the groups.

In this paper, we use the fuzzy C-means algorithm and interpret the degree of membership of a fund to the wrong cluster as the *degree of inaccuracy* of its name. For example, if a fund name includes the word *small* but belongs to the group *large* with a degree of membership of 0.7, we conclude that this fund name has a degree of inaccuracy of 0.7. Similarly, if a fund name includes the word *value* and belongs to the group *growth* with a degree of membership of 0.1, we conclude that this fund name has a degree of inaccuracy of 0.1.

To determine the name of each group, we look at the names of funds that belong to each group with a degree of membership of at least 0.5, as these funds are more likely to belong to this group than to the other one. Suppose the majority of funds in a group has a name including the word *small* (vs. *large*), this group is labelled *small*.[9]

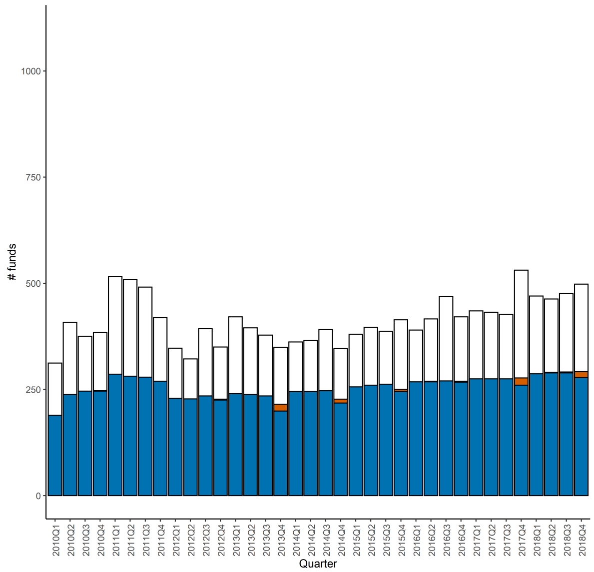
Finally, based on this degree of inaccuracy, we also create a dummy variable *inaccurate*. When a fund's degree of membership to the wrong group is higher than or equal to 0.5, we identify this fund's name as inaccurate. In the example above, the fund featuring the term *small* in its name is inaccurate, while the fund featuring the word *value* in its name is accurate.

Figure 1 illustrates the results of the cluster analysis. First, the classification of each cluster is straightforward as each cluster is dominated by funds having a specific term in their name (either *small* or *large* for the cluster on the SMB dimension determined using $\beta_{S,i}$, and either *growth* or *value* for the cluster on the HML dimension determined using $\beta_{H,i}$). Second, there is some heterogeneity across the dimensions and also across time. For instance, a larger fraction of HML funds is classified as inaccurate compared to SMB funds.

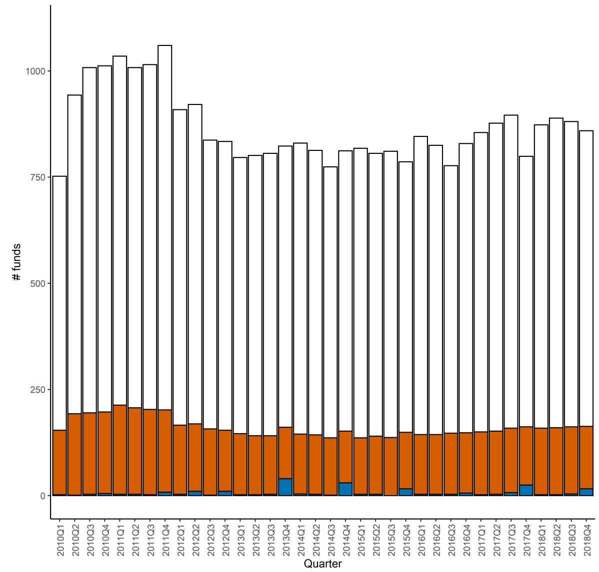
More detailed information on the proportion of inaccurate fund names, per year, per type of fund name, and per dimension on which they are inaccurate is reported in Table 3.[10] On average, each year, around 25% of funds in our sample are inaccurate in at least one dimension. This proportion slightly varies over time but always remains above 18%. The phenomenon of inaccurate naming, therefore, tends to be widespread. Looking at the specific

[9]When using a clustering algorithm as fuzzy C-means or K-means, one should know that the different initial partitions might lead to different cluster results. However, this is not a concern for us. Indeed, as suggested in Jain (2010), we did a robustness analysis and run the clustering algorithm with several initial partitions: the *inaccurate* dummy variable remains the same.

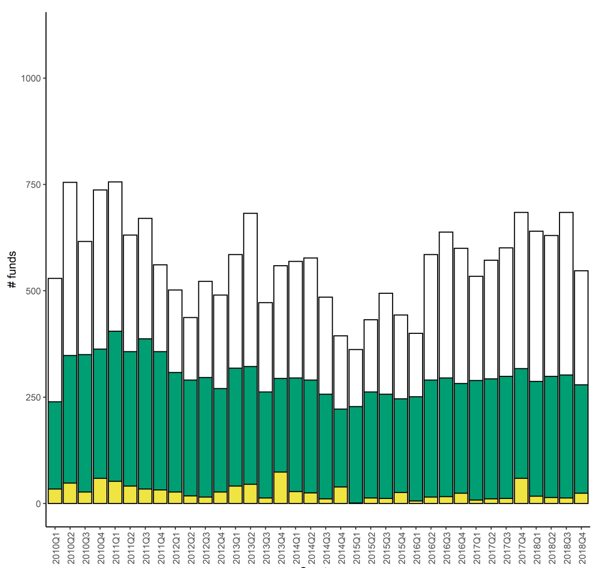
[10]When a fund name contains both a term referring to the SMB dimension (*small* or *large*) and a term referring to the HML dimension (*value* or *growth*), it can be found to be inaccurate either on the SMB dimension, or on the HML dimension, or on both dimensions. Hence, there are in total 8 possible accurate/inaccurate name cases. Table B displays the various cases.



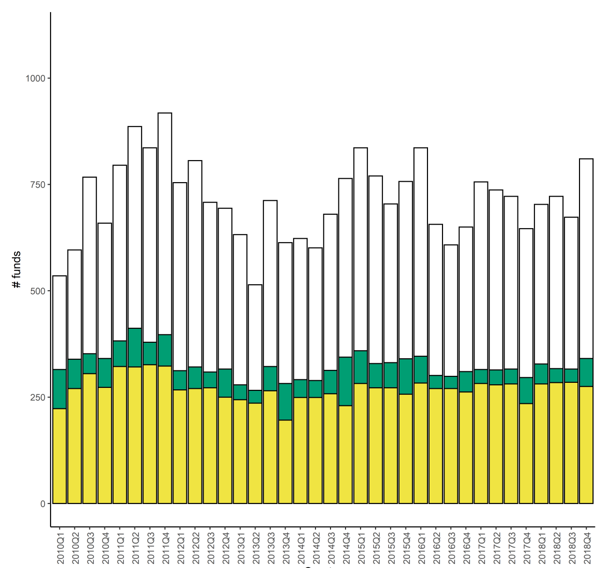
(a) **Panel A: Cluster *Small***



(b) **Panel B: Cluster *Large***



(c) **Panel C: Cluster *Growth***



(d) **Panel D: Cluster *Value***

Fig. 1. **Number of funds per cluster per year quarter**

Figure 1 shows the number of funds per cluster per year quarter. In Panels A and B, funds whose name contain the word *small* are in blue and funds with the word *large* in their name are in red. Other fund names are in white. In Panels C and D, fund names referring to *growth* are in green and funds having the word *value* in their name are in yellow. Other names are in white. In total, the sample consists of 2,669 fund names between 2010 and 2018.

dimensions of inaccuracy, we see that a much bigger proportion of funds is inaccurate on the HML dimension than on the SMB dimension. This finding is not surprising given that, according to the SEC Names Rule (35d-1), names referring to the SMB dimension must be used in a "reasonable manner", while there is no such clear restriction for names referring to the HML dimension (SEC, 2001). The findings, therefore, suggest that a Names Rule is effective. Finally, we find that only a very small proportion of funds is inaccurate on both dimensions.

Focusing on the results per type of name, Table 3 reveals that the funds that most often feature an inaccurate name are those referring to the two terms *small & growth* in their name. More than 30% of these fund names are indeed found to be inaccurate every year. On the other end of the spectrum, the funds featuring the term *large* are very rarely classified as being inaccurate.

In addition, we also provide descriptive statistics about the prevalence of this practice of inaccurate naming per fund, ie is a fund inaccurate just once or twice in its life-cycle or does it occur more frequently. Figure 2 sheds light on this issue and shows that a large fraction of funds indeed provides inaccurate information for just a limited number of times, but an important fraction of funds do exhibit such misnaming practice for many periods; with some funds even having an inaccurate name for as long as 8 years (almost the entire sample period). Such high frequencies hint at a rather deliberative process by the fund managers.

Finally, we also provide descriptive evidence on the distribution of the inaccuracies over the quarters in a year. In line with the tournament hypothesis that is analyzed in the next section, Table 4 indeed confirms that most inaccuracies occur during the fourth quarter of a given year. This non-random pattern suggests that the practice is intentional.

Table 3: **Inaccurate mutual fund names**

Table 3 shows the proportion (in %) of fund names which are found to be inaccurate each year, per type of name, and per dimension on which they are found to be inaccurate. SMB refers to inaccurate fund names associated with the SMB dimension (cases (2) and (6) in Table B). HML refers to inaccurate fund names associated with the HML dimension (cases (4) and (7)). SMB & HML refers to inaccurate fund names associated with both dimensions (case (8)). Note that the number in All is not necessarily the sum of the numbers in SMB, HML, and SMB & HML. Indeed, a fund name related to two dimensions (*small & value* for example), can be inaccurate on the different dimensions during the year. For example, in a given quarter of a year, the use of *small* is inaccurate while the use of *value* is accurate, and on another quarter of the same year the use of both *small* and *value* is inaccurate. So, this fund name is included in once on row All, as well as once on row SMB and once on row SMB & HML.

Name	2010	2011	2012	2013	2014	2015	2016	2017	2018
All funds									
All	28.56	22.09	20.75	36.06	31.70	25.51	18.12	23.24	20.75
SMB	0.93	0.89	2.61	5.71	3.96	2.56	1.65	5.36	3.89
HML	27.73	21.40	18.51	29.18	26.68	22.31	16.35	17.52	16.75
SMB & HML	0.00	0.20	0.12	2.08	1.45	0.90	0.13	0.72	0.47
Small	3.28	3.82	9.91	15.79	12.07	5.43	3.82	12.50	9.52
Small & Growth	75.71	56.00	65.00	69.49	60.66	57.14	36.51	31.34	35.71
Small & Value	10.71	4.30	7.89	37.18	23.75	16.47	9.09	19.77	9.30
Large	0.00	0.00	0.00	7.89	0.00	2.38	0.00	9.43	10.00
Large & Growth	3.80	6.02	7.94	26.79	25.45	24.49	12.50	16.00	17.31
Large & Value	22.08	14.10	11.11	36.84	26.92	17.65	16.67	23.64	19.64
Growth	41.76	31.67	27.83	36.84	46.08	46.60	33.17	29.03	32.08
Value	37.93	31.63	22.99	46.25	31.79	15.48	16.13	29.94	21.21

4. Reasons for inaccurate fund names: the tournament hypothesis

In this section, we focus on the reasons for the observed deviations in the name dimension as motivated by the tournament hypothesis of Brown et al. (1996) and Chevalier and Ellison (1997). More precisely, five elements are combined to develop this hypothesis. First, the mutual fund industry can be seen as a tournament in which funds with similar investment

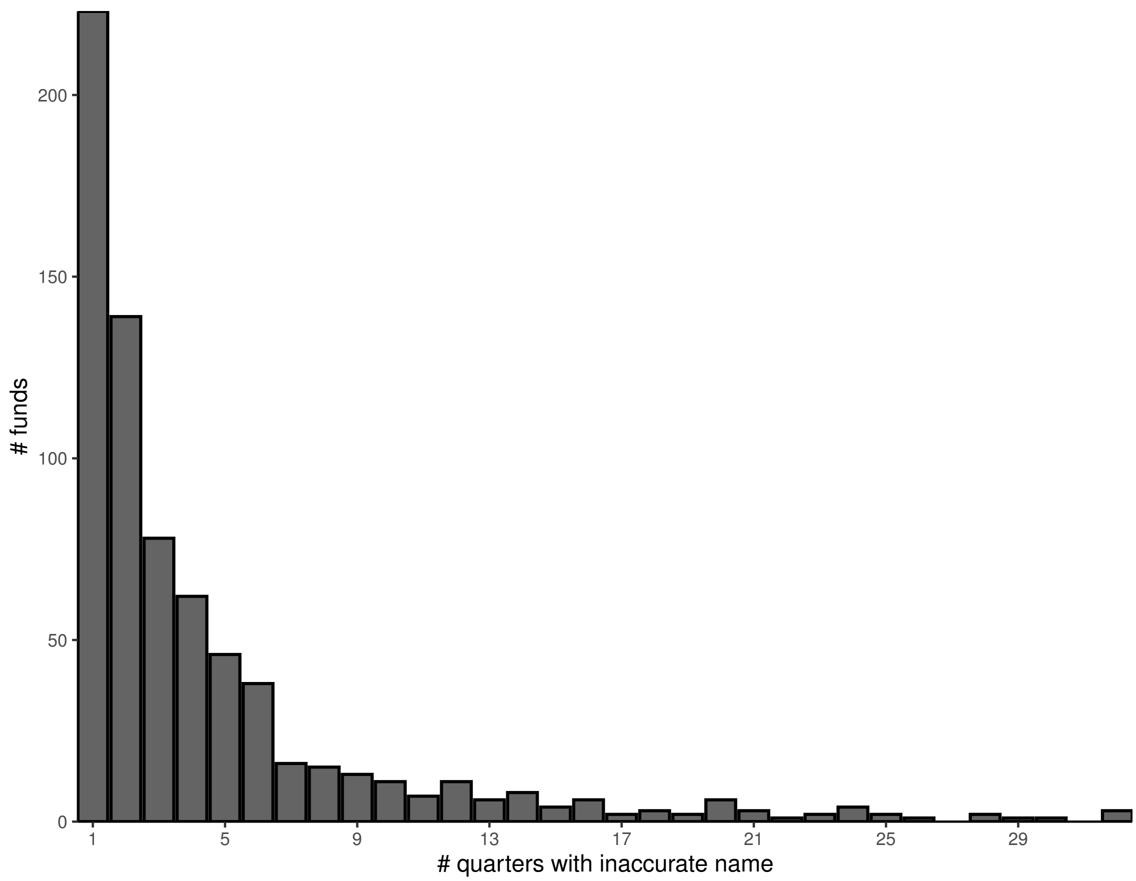


Fig. 2. **Frequency of inaccurate fund names**

Figure 2 shows the number of funds having an inaccurate name for a given number of quarters. In total, the sample consists of 3,008 inaccurate fund-quarter observations between 2010 and 2018.

Table 4: **Quarter of deviation**

Table 4 reports the deviations (in %) from the investment style suggested by the name (fund names switching from being accurate to being inaccurate) per quarter and per dimension (SMB, HML, or both SMB & HML).

	Q1	Q2	Q3	Q4
All	14	17	15	53
SMB	6	12	7	75
HML	16	19	17	48
SMB & HML	0	0	0	100

styles compete with each other.[11] The funds at the top of the ranking within a style group win the tournament, while those at the bottom lose it. Second, the vast majority of these rankings are produced at the end of a calendar year, based on funds' performance over the year (Brown et al., 1996). Third, according to the tournament hypothesis, the funds appearing at the top of these end-of-year rankings, ie the winners of the tournament, receive higher fund inflows as compared to the losers (Brown et al., 1996). Fourth, in contrast to the fund inflows for the winners, there are no significant fund outflows for the loser. Hence, there is empirical evidence for an asymmetric return-flow relationship (Brown et al., 1996; Chevalier and Ellison, 1997). While funds with better performance are rewarded with higher inflows of money, worse-performing funds do not necessarily suffer from outflows. Fifth, managers' compensation is determined based on (new) assets under management. Therefore, their compensation is positively related to the amount of money flowing into the fund (Brown et al., 1996), giving them strong incentives to be among the winners at the end of the year.

Assembling these puzzle pieces results in the following hypothesis. As managers' compensation is linked to fund's inflows, which, in turn, is the consequence of better performance, managers have incentives to be among the end-of-year winners. For this reason, when their performance over the course of a year is worse than their competitors', they have an incentive to deviate from the investment style stated in their name in an attempt to catch up with competitors by the end of the year, when rankings are finalized. This behavior can be explained with prospect theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992). Due to the convexity of the utility function in the loss area, individuals are willing to take more risk. If this deviation is successful, their funds will receive higher new fund inflows, and their compensation will increase. However, given the asymmetric return-flow relationship, if this deviation is not successful and leads to worse end-of-year performance, managers would, on the other hand, not necessarily be penalized by investment outflows,

[11] This tournament characteristic is, for instance, illustrated by the periodic performance ranking of funds by business magazines and financial service firms.

and their compensation would not be drastically affected.

In the context of the name inaccuracy, the tournament hypothesis can be supported by three different pieces of evidence. First, if the tournament hypothesis holds, funds would deviate more often at the end of the year. Hence, we expect fund name deviations to happen mainly during the fourth quarter of a year. This hypothesis is motivated by the attempt to catch up with the winning funds by the end of the year when performance rankings are calculated. Second, the rank of a fund's performance in comparison to its competitors would determine the degree to which a fund deviates from the stated investment style, ie the degree of inaccuracy: the lower the rank, the higher the incentive to "gamble" and to deviate from the stated investment style. Third, as the tournament hypothesis implies that the decision to deviate is deliberate and is part of a strategy, we expect a difference between funds whose names are only inaccurate for one quarter (which could be due to inattention) and funds whose names are inaccurate for multiple quarters. A repetitive character would indeed point at a deliberate strategy used by managers rather than inattention. In particular, the performance ranking would determine the degree of inaccuracy of funds whose names are inaccurate for at least two quarters. This would not be the case for funds featuring an inaccurate name for only one quarter.

Hence, in what follows, we provide evidence for the support of the tournament hypothesis. First, we look at the timing of the inaccuracy to identify potential patterns. Second, we analyze the determinants of the degree of inaccuracy and investigate whether performance ranking is indeed a significant determinant of a funds inaccuracy. Finally, we distinguish funds deviating for only one quarter from funds deviating for at least two quarters and conduct the previous analysis on each sample separately.

4.1. Point in time of the inaccuracy

As already reported, Table 4 reports the percentage of funds deviating in each quarter.[12] The results show that most deviations (53%) occur during the last quarter of a year, while the rest is equally distributed among the other three quarters. Accordingly, most fund name deviations come up at the end of the year, when relative performance rankings are created. Thus, this finding is in line with the tournament hypothesis, which predicts that funds would deviate from their stated investment strategy in an attempt to increase their performance and catch up with best-ranked competitors by the end of the year when performance rankings are issued. Interestingly, this result holds for all dimensions but amplifies in the more strictly regulated dimensions (SMB or SMB & HML), in which we observe, in general, fewer deviations (see Table 3).

4.2. Determinants of inaccuracy and fund's performance rank

Next, we investigate which determinants explain the degree of inaccuracy of mutual fund names, ie the degree to which a fund deviates from the investment styles stated in its fund name. In particular, we are interested in whether the rank of a fund based on its performance during the first three quarters of a year determines the degree to which this fund deviates during the last quarter of the year, as suggested by the tournament hypothesis.

To this end, we estimate a panel regression in which the dependent variable is the degree of inaccuracy in the fourth quarter of a calendar year. We relate this inaccuracy to various fund characteristics observed in the first three quarters of the year such as fund flows, fund age, costs, and risk. The rank of a fund is determined based on its performance with respect to its competitors, defined as the funds having the same name; ie a fund named *small* competes with all the other funds named *small*. And the performance is proxied by the fund return over the first three quarters of a year. Finally, we control for general market conditions by including quarter fixed effects and investment style fixed effects. The samples

[12] To make results across categories comparable, we report relative and not absolute numbers.

of funds in this analysis are the one defined in Table C with a particular focus on the last quarter of a year.Hence, we estimate the following panel regression model:

$$
\begin{aligned}
\text{Inaccuracy}_{f,t} =& \gamma_0 + \gamma_1 \text{Rank}_{f,[t-3,t-1]} + \gamma_2 \text{Flow}_{f,[t-3,t-1]} + \gamma_4 \text{Risk}_{f,[t-3,t-1]} \\
&+ \gamma_5 \text{Fund Size}_{f,[t-3,t-1]} + \gamma_6 \log(\text{Age})_{f,[t-1]} + \gamma_6 \text{Size}_{f,[t-3,t-1]} \\
&+ \gamma_7 \text{Expense Ratio}_{f,[t-7,t-4]} + \gamma_8 \text{Turnover Ratio}_{f,[t-7,t-4]} \\
&+ \text{Quarter dummy}_t + \text{Style dummy}_f + \epsilon_{f,t-1},
\end{aligned} \tag{3}
$$

where $\text{Inaccuracy}_{f,t}$ is the fund degree of inaccuracy during a given fourth quarter, Rank is the fund rank calculated using the performance over the first three quarters of the year, Flow is the average fund flow over the first three quarters of the year, Risk is the average fund risk over the first three quarters of the year, log(Age) is the fund age at the end of the first three quarters of the year, Size is the average fund size over the first three quarters of the year, Expense Ratio is the average expense ratio over the previous year and Turnover Ratio is the average turnover ratio over the previous year. Quarter fixed effects are used, as well as style fixed effects where style is defined as in the second panel of Table 3. The standard errors are clustered at the fund level.

The coefficient of interest with respect to the tournament hypothesis, γ_1, captures the marginal effect of the fund's respective rank on the degree of inaccuracy and indicates whether its relative performance drives the degree of a fund's inaccuracy. The results for our sample of funds that have a name referring to either the size (SMB), the investment strategy (HML) dimension, or both dimensions at the same time are reported in Table 5.

Results reported in Table 5 support our hypothesis: the lower the rank of a fund's performance with respect to its competitors, the higher the degree to which the fund deviates from the investment style stated in its name. Thus, this finding suggests that funds that are relative losers change their investment styles, which results in name inaccuracy, to improve their relative performance. In columns (2)–(4) of Table 5, we separate the analysis between

Table 5: **Determinants of Inaccuracy of Fund Names End-of-Year**

Table 5 reports the result of a panel regression investigating the determinants of the degree of inaccuracy of a fund (ie the degree to which a fund deviates from the investment style stated in its name) during the last quarter of a year. Quarter fixed effects and style fixed effects are used. The dependent variable *Inaccuracy* is between 0 and 100. The independent variables are all standardized. Column (1) reports the results for the whole sample of funds, column (2) reports the results for funds with a name related to the SMB dimension, column (3) reports the results for funds with a name related to the HML dimension, and column (4) reports the results for funds with a name related to both the SMB and HML dimensions. The standard errors are clustered at the fund level and reported in parentheses. * , **, and *** represent statistical significance at the 10%, 5%, and 1% level, respectively.

	(1) All	(2) SMB	(3) HML	(4) SMB & HML
Rank	-1.4075***	-0.6897**	-1.7154***	-0.1438
	(0.3516)	(0.3392)	(0.4088)	(0.3859)
Flow	-2.6502***	-2.4319***	-1.7844***	-1.7471***
	(0.3395)	(0.3217)	(0.3995)	(0.3714)
Risk	0.7788*	1.2884**	0.3524	0.5233
	(0.4370)	(0.5375)	(0.4650)	(0.5123)
log(Age)	-0.0495	-0.0769	-0.1407	-0.0570
	(0.4247)	(0.3277)	(0.5091)	(0.4737)
Size	-0.6636	0.7399*	-1.0468*	0.7343
	(0.5187)	(0.4339)	(0.6329)	(0.5553)
Expense Ratio	0.5825	-0.0163	1.0555*	0.0025
	(0.4274)	(0.2236)	(0.6139)	(0.4216)
Turnover Ratio	-0.6028*	0.2679	-0.9272**	0.4124
	(0.3592)	(0.3646)	(0.3934)	(0.6689)
Quarter Fixed Effect	Y	Y	Y	Y
Style Fixed Effect	Y	Y	Y	Y
Observations	5,936	3,016	4,511	1,667
Adj. R^2	0.1076	0.0826	0.0747	0.0575

funds that provide a size indication in their name (2), those that indicate an investment strategy (3), and those that refer to both dimensions (4). The main results are similar to the regression that includes all funds.

In addition to the performance ranking, other variables are found to be significant determinants. One of them is fund risk. When a fund has higher idiosyncratic risk, it deviates more from the stated investment style.

Moreover, we find a significant effect of fund flows on the degree of inaccuracy. The fund flows in the previous three quarters are significantly negatively related to the degree to which a fund deviates from its suggested investment strategy in the name.

We also include fund size, age, expense ratio and turnover ratio to analyze whether larger, older, more expensive, and more active funds have a higher probability of providing an inaccurate fund name but do not find any systematic significant impact.

4.3. Inattention vs. deliberate strategy

While the results in Table 5 support our hypothesis and provide evidence that managers' decisions to deviate from the strategy are linked to a relatively bad rank in performance in the previous quarters, we cannot rule out that a deviation can also happen by inattention. Hence, in the next step, we test whether this decision is based on a manager's inattention or whether it is because of a deliberate strategy.

Figure 2 in Section 3 shows that a large fraction of funds provides only for a very short time an inaccurate fund name, while other funds do so for a long time. Hence, we test whether funds that differ in the frequency of inaccurate information also differ with respect to the motivation. The implicit assumption behind this analysis is that short-term inaccuracy represents an inattention motive while a deliberate strategy would drive repeated inaccuracy. Therefore, we use the repetitive nature of a name inaccuracy as an indication that the decision to deviate is part of the fund's strategy to achieve a higher rank in the last quarter of the year.

We perform a similar analysis as in Table 5 but we include a dummy variable *Deliberate*, which is equal to 1 when a fund has an inaccurate name at least twice at end-year and 0 otherwise. We further interact this *Deliberate* dummy with the independent variables. A negative and significant interaction term *Rank x Deliberate* would provide evidence that deviating from the style mentioned in the name is used a strategy to try to improve a fund's rank by the end of the year, when the performance over the first three quarters of a year was bad relative to competitors. Therefore, a negative and significant *Rank x Deliberate* would support the tournament hypothesis.

The results of Table 6 support the tournament hypothesis as the intersection term *Rank x Deliberate* drives inaccurate mutual fund names. This is true for all funds, and this is especially true for the funds deviating on the HML dimension (with a name including the term *growth* or *value*). This is interesting given that this is the name dimension that is less strictly regulated by the SEC. Therefore, in this dimension, funds can deviate from the investment style stated in their name while still complying with the Names Rule (which would not especially be the case if they deviate on the SMB dimension).

Therefore, we conclude that the results of our analyses confirm that managers use deviations from the investment styles stated in the name (leading to an inaccurate name) as a strategy when their performance over the course of the year is worse than their competitors, to try to achieve a higher performance rank by the end of the year, when rankings are issued.

5. Consequences of inaccurate names

Finally, we analyze the possible consequences of deviating from the investment style stated in the name. In particular, we are interested in central fund characteristics before and after a name becomes inaccurate. Our goal is to investigate whether implementing a deviation strategy to achieve a higher performance ranking is effective and to understand the investors' reaction to an inaccurate name.

Table 6: **Tournament hypothesis**

Table 6 reports the result of a panel regression to investigate the determinants of the degree of inaccuracy of a fund (ie the degree to which a fund deviates from the investment style stated in its name) in the last quarter of a year. Quarter fixed effects and style fixed effects are used. The dependent variable *Inaccuracy* is between 0 and 100. The variables are all standardized. Column (1) reports the results for the whole sample of funds, column (2) reports the results for funds with a name related to the SMB dimension, column (3) reports the results for funds with a name related to the HML dimension, and column (4) reports the results for funds with a name related to both the SMB and HML dimensions. The standard errors are clustered at the fund level and reported in parentheses. * , **, and *** represent statistical significance at the 10%, 5%, and 1% level, respectively.

	(1) All	(2) SMB	(3) HML	(4) SMB & HML
Rank	-0.6081**	-0.7894***	-0.5677	-0.1491
	(0.2975)	(0.3057)	(0.3540)	(0.3584)
Flow	-1.8283***	-1.7955***	-1.1011***	-0.9811***
	(0.2558)	(0.2608)	(0.3181)	(0.3447)
Risk	0.8911**	0.8513*	0.6482*	0.5393
	(0.3495)	(0.4657)	(0.3923)	(0.5529)
log(Age)	-0.3461	-0.2157	-0.4187	-0.4218
	(0.3499)	(0.2943)	(0.4591)	(0.4686)
Size	-0.6184*	0.3058	-0.8028	0.2776
	(0.3701)	(0.3503)	(0.4947)	(0.4799)
Expense Ratio	0.6048	-0.0541	1.1201*	-0.5824*
	(0.4485)	(0.1952)	(0.6215)	(0.3383)
Turnover Ratio	0.0893	0.3139	-0.2365	0.3613
	(0.3230)	(0.3128)	(0.3179)	(0.7432)
Deliberate	20.9998***	10.6864***	18.8223***	5.9938***
	(0.9022)	(1.3636)	(0.9686)	(1.1550)
Rank x Deliberate	-1.8972**	0.9680	-2.7918***	-0.1655
	(0.9534)	(1.0306)	(1.0284)	(0.9721)
Flow x Deliberate	-1.9527*	-2.6235**	-1.4342	-2.4821***
	(1.0302)	(1.1797)	(1.0880)	(0.9243)
Risk x Deliberate	-1.6300	0.5488	-1.3207	-0.3310
	(1.0438)	(1.1256)	(1.1255)	(1.0464)
log(Age) x Deliberate	-0.6494	-0.9827	-0.4947	0.9381
	(1.0050)	(1.3612)	(1.0818)	(1.2674)
Size x Deliberate	-1.1070	0.4880	-1.0951	1.1590
	(1.3405)	(1.7987)	(1.5064)	(1.5567)
Expense Ratio x Deliberate	1.3837	-3.4188	1.1345	2.1047
	(3.1122)	(3.0261)	(2.6723)	(1.3825)
Turnover Ratio x Deliberate	-1.6656	0.7388	-1.3827	0.9526
	(1.3261)	(0.8392)	(1.5209)	(1.1797)
Quarter Fixed Effect	Y	Y	Y	Y
Style Fixed Effect	Y	Y	Y	Y
Observations	5,936	3,016	4,511	1.667
Adj. R^2	0.2106	0.1320	0.1710	0.0850

Given that deviating from the stated investment style is not random, we follow Cooper et al. (2005) and rely on propensity score matching. This approach allows us to match funds similar in every characteristic but the naming accuracy. To obtain the propensity score, we estimate a logit regression using the same explanatory variables as in Table 5. Moreover, to ensure that we match funds that are as similar as possible and to control for unobservable characteristics across time, we restrict ourselves to matches of funds within a year-quarter-investment style category. Once a match is found, we compute *abnormal* fund characteristics as the difference between the characteristics of the treated fund (whose name switches from accurate to inaccurate) and of the control fund (whose name keeps being accurate) when the name becomes inaccurate.

The results are reported in Table 7. They are reported for each dimension on which a fund name can be inaccurate (only SMB, only HML, or both). Moreover, given the results of the previous section highlighting that funds having an inaccurate name for one quarter are not the same as funds having an inaccurate name for more than one quarter, we report three different results: (1) using the full sample of funds having an inaccurate name during the last quarter of a year ($N >= 1$), (2) using a sample of funds whose names are only inaccurate once during the last quarter of a year ($N = 1$), (3) and using a sample of funds whose names are inaccurate more than once during the last quarter of a year ($N > 1$).

Our first characteristic of interest is fund flows, as those indicate whether and how investors react to such name inaccuracy. An increase in fund flows could be interpreted as an increased interest in the fund from investors. The increased fund flow also means that managers receive higher compensation (managers' compensation is increasing with the fund's size). On the other hand, a decrease in fund flows could indicate that investors take note of funds' deviating behavior and decide to invest in alternative products.

Overall, Panel A in Table 7 shows that funds do experience statistically significant abnormal outflows. This result, therefore, suggests that investors do respond to inaccurate information. While the ultimate goal of the deviation strategy implemented by fund managers is to attract new investments into the funds in order to increase their compensation, the results show that this is not effective, on average. There are, however, again differences across dimensions. While funds that refer to the SMB dimension in the name experience statistically significant lower fund flows than their counterparts, funds referring to the HML dimension experience lower fund flows, which, however, are not always statistically significant and never significant at a 5% significance level.

Second, we examine potential consequences in a funds' return rank. If the deviation from the stated investment strategy leads to a higher rank, this indicates that the strategy to use deviations to increase the fund rank is successful. However, we do not find evidence that such deviating behaviour consistently pays off in terms of improving the fund's performance rank. The strategy sometimes even negatively impacts the rank. In this case, not deviating from the name-suggested style would have delivered better relative performance.

Third, we investigate the consequence in terms of risk. A higher risk would highlight that the funds "gamble" by deviating to try to achieve a better ranking. Moreover, if risk increases and the rank is not improved, investors do not benefit from a fund's decision to deviate from its stated investment style. Bearing higher risk but not benefiting in terms of returns is clearly undesirable for investors.

The results of abnormal risk reported in Table 7 show that risk significantly increases. This result, taken together with the abnormal rank results, shows that a fund's risk-return trade-off is worse after a deviation. Moreover, we observe a significant difference between funds that deviate on the SMB dimension and those who deviate on the HML dimension: the trade-off is even worse for those deviating on the more strictly regulated SMB dimension.

In general, the results reported in Table 7 highlight that funds perform very poorly after deviating from the investment style stated in their name. This suggests that these funds gambled but lose. Hence, our findings are in line with the tournament hypothesis.

Moreover, as highlighted by the abnormal flows, investors are only to some extent able to notice this deviation, at least not in a timely manner. This adds an additional layer

Table 7: **Consequence of inaccurate names**

Table 7 reports the consequence of having an inaccurate name. A propensity score matching technique is used to find a fund with an accurate name which is the most closely related to a fund with an inaccurate name. The propensity score are obtained with a logit regression using as dependent variable a dummy that is equal to 1 when the fund name becomes inaccurate and 0 when the fund name remains accurate. The explanatory variables are the same as in Table **??**. The control fund is also required to be in the same year-quarter-investment style category as the treated fund. The abnormal characteristic is the difference between the characteristic of the treated fund and the characteristic of the control fund. Panel A reports the abnormal characteristics. Panel B reports the t-statistic for the difference in abnormal characteristics across the dimensions in the fund name (SMB, HML, or both). The results are also split for different samples: $N >= 1$ is a sample of funds whose names are inaccurate at least once during the last quarter of a year, $N = 1$ is the sample of funds whose names are inaccurate only once, and $N > 1$ is the sample of funds whose names are inaccurate for more than once. * , **, and *** represent statistical significance at the 10%, 5%, and 1% level, respectively.

	(1) Flow	(2) Rank	(3) Risk
Panel A: Abnormal characteristics			
All			
(1a) $N >= 1$	-1.5497***	0.0023	0.5588***
(1b) $N = 1$	-1.6483*	0.0361	0.4596***
(1c) $N > 1$	-1.5037**	-0.0134	0.6051***
SMB			
(2a) $N >= 1$	-2.5017**	-0.0165	0.9397***
(2b) $N = 1$	-2.1336	0.0273	0.9048***
(2c) $N > 1$	-2.7352*	-0.0443	0.9618***
HML			
(3a) $N >= 1$	-1.1238*	0.0112	0.4275***
(3b) $N = 1$	-1.3175	0.0481*	0.3003***
(3c) $N > 1$	-1.0385	-0.0051	0.4836***
SMB & HML			
(4a) $N >= 1$	-4.6681**	-0.0657	1.1707***
(4b) $N = 1$	-5.4133	-0.1586**	0.8859***
(4c) $N > 1$	-4.4368*	-0.0368	1.2591***
Panel B: Differences in characteristics			
SMB vs. HML			
(2a) - (3a)	-1.0976	-0.7451	4.0381***
(2b) - (3b)	-0.4174	-0.3637	3.4502***
(2c) - (3c)	-1.0254	-0.8028	2.7175***
SMB & HML vs. SMB			
(4a) - (2a)	-0.9633	-0.6288	0.9553
(4b) - (2b)	-0.7823	-2.2825**	-0.0641
(4c) - (2c)	-0.6195	0.0746	0.9486
SMB & HML vs. HML			
(4a) - (3a)	-1.6943*	-1.0653	3.4186***
(4b) - (3b)	-0.9999	-2.9577**	2.1881*
(4c) - (3c)	-1.3785	-0.3452	2.8451***

of concern: investors are worse off but experience difficulties to react. Thus, our results highlight the need for stricter regulations, not only when a fund changes its name or is created but also over its whole life cycle.

6. Conclusion

A significant fraction of mutual funds refers to an investment style in their name. This paper explores whether these funds stick to these highlighted investment styles or deviate from it, resulting in inaccurate fund names. In particular, we want to understand which funds provide inaccurate name information, why they feature an inaccurate name and the consequences of such inaccuracy.

To answer these research questions, we construct a fund name history dataset based on fund prospectuses in the EDGAR database of the SEC. Extracting information from more than 400,000 prospectuses between January 2010 and December 2018 allows us to build a detailed name history dataset for US equity mutual funds. To identify inaccurate fund names, we then use a cluster analysis technique.

In the first part of the paper, we document that a significant fraction of mutual funds is associated with an inaccurate fund name. We find that 33% of US equity mutual funds have an inaccurate name at least once in their life-cycle. Second, we focus on two specific dimensions on which a name can be inaccurate: the size dimension (name including the terms *small* or *large*), and the investment strategy dimension (name including the terms *growth* or *value*). The size dimension displays relatively few inaccurate fund names, while inaccurate fund names are abundant in the investment strategy dimension. This result is not surprising given that the SEC regulates more strictly the names related to the size dimension (see Rule 35d-1) than those related to the investment strategy dimension.

Regarding the characteristics of funds with inaccurate fund names, we find that funds that deviate from the investment style stated in their names experience lower fund inflows and are riskier. More importantly, we find that funds that are relative losers in performance compared to their competitors deviate more from the stated investment styles in the names. This finding can be rationalized by prospect theory and supports our tournament hypothesis, which states that funds deliberately use deviations from the style in the name as a strategy to climb the performance ranking. This hypothesis is further supported by the fact that this effect is more pronounced when funds deviate from the investment style stated in their name in a repeated manner (which hints at the use of the deviation as a strategy rather than mere inattention).

Finally, to shed light on the consequences of inaccurate fund names, we estimate potential abnormal characteristics of inaccurate funds with respect to funds that do not deviate from the investment style included in their name. We rely on propensity score matching to overcome potential endogeneity problems. By looking at funds' performance ranking and idiosyncratic risk, respectively, we find that funds do not improve their ranking while being exposed to higher idiosyncratic risk after deviating from the stated investment style. This consequence is even more pronounced for funds whose name is more strictly regulated by the SEC. However, investors have difficulties in responding to these disadvantageous characteristics of funds having an inaccurate name, at least in a timely manner.

Overall, our results highlight the role of transparency and a precise regulation in the fund names.

References

Abis, S., 2017. Man vs. machine: Quantitative and discretionary equity management. Working Paper .

Baghai, R., Becker, B., Pitschner, S., 2019. The private use of credit ratings: Evidence from investment mandates. Swedish House of Finance Research Paper No. 18-13 .

Bams, D., Otten, R., Ramezanifar, E., 2017. Investment style misclassification and mutual fund performance. 28th Australasian Finance and Banking Conference. Available at SSRN: https://ssrn.com/abstract=2648375 .

Barber, B. M., Odean, T., Zheng, L., 2005. Out of sight, out of mind: The effects of expenses on mutual fund flows. The Journal of Business 78, 2095–2120.

Bezdek, J. C., 1981. Pattern recognition with fuzzy objective function algorithms. Advanced applications in pattern recognition, Plenum, New York (N.Y.).

Brown, K. C., Harlow, W. V., Starks, L. T., 1996. Of tournaments and temptations: An analysis of managerial incentives in the mutual fund industry. The Journal of Finance 51, 85–110.

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

Carhart, M. M., 1997. On Persistence in Mutual Fund Performance. Journal of Finance 52, 57–82.

Chevalier, J., Ellison, G., 1997. Risk taking by mutual funds as a response to incentives. Journal of Political Economy 105, 1167–1200.

Choi, D., Kahraman, B., Mukherjee, A., 2016. Learning about mutual fund managers. The Journal of Finance 71, 2809–2860.

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

Cremers, K. J. M., Petajisto, A., 2009. How Active Is Your Fund Manager? A New Measure That Predicts Performance. The Review of Financial Studies 22, 3329–3365.

DiBartolomeo, D., Witkowski, E., 1997. Mutual fund misclassification: Evidence based on style analysis. Financial Analysts Journal 53, 32–43.

Dunn, J. C., 1973. A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters. Journal of Cybernetics 3, 32–57.

Espenlaub, S., ul Haq, I., Khurshed, A., 2017. It's all in the name: Mutual fund name changes after sec rule 35d-1. Journal of Banking & Finance 84, 123–134.

Gallaher, S. T., Kaniel, R., Starks, L. T., 2015. Advertising and mutual funds: From families to individual funds. CEPR Discussion Paper No. DP10329 .

Hillert, A., Niessen-Ruenzi, A., Ruenzi, S., 2016. Mutual fund shareholder letter tone-do investors listen? Working Paper .

Huang, J., Sialm, C., Zhang, H., 2011. Risk shifting and mutual fund performance. The Review of Financial Studies 24, 2575–2616.

Investment Company Institute, 2019. What us households consider when they select mutual funds, 2018. ICI Release, Available at https://ici.org/pdf/per25-03.pdf .

Jain, A. K., 2010. Data clustering: 50 years beyond K-means. Pattern Recognition Letters 31, 651–666.

Jain, P. C., Wu, J. S., 2000. Truth in mutual fund advertising: Evidence on future performance and fund flows. The Journal of Finance 55, 937–958.

Kacperczyk, M., Sialm, C., Zheng, L., 2008. Unobserved actions of mutual funds. The Review of Financial Studies 21, 2379–2416.

Kahneman, D., Tversky, A., 1979. Prospect theory: An analysis of decisions under risk. Econometrica 47, 263–291.

Kim, M., Shukla, R., Tomas, M., 2000. Mutual fund objective misclassification. Journal of Economics and Business 52, 309–323.

Kim, T.-H., White, H., Stone, D., 2005. Asymptotic and bayesian confidence intervals for sharpe-style weights. Journal of Financial Econometrics 3, 315–343.

Kostovetsky, L., Warner, J. B., 2020. Measuring innovation and product differentiation: Evidence from mutual funds. The Journal of Finance 75, 779–823.

Krakow, J., Schäfer, T., 2020. Mutual funds and risk disclosure: Information content of fund prospectuses. Working Paper .

Kumar, A., Niessen-Ruenzi, A., Spalt, O. G., 2015. What's in a name? mutual fund flows when managers have foreign-sounding names. The Review of Financial Studies 28, 2281–2321.

Mason, A., Mcgroarty, F., Thomas, S., 2012. Style analysis for diversified US equity funds. Journal of Asset Management 13, 170–185.

Mateus, I., Mateus, C., Todorovic, N., 2019. Benchmark-adjusted performance of US equity mutual funds and the issue of prospectus benchmarks. Journal of Asset Management 20, 15–30.

SEC, 2001. Investment Company Act Release No. 24828 ("Names Rule Adopting Release"). Available at https://www.sec.gov/rules/final/ic-24828.htm .

SEC, 2020. Press Release: SEC Requests Comment on Fund Names Rule). Available at https://www.sec.gov/news/press-release/2020-50 .

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

Sirri, E. R., Tufano, P., 1998. Costly search and mutual fund flows. The Journal of Finance 53, 1589–1622.

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

Thaler, R. H., 2016. Behavioral economics: Past, present, and future. American Economic Review 106, 1577–1600.

Tversky, A., Kahneman, D., 1992. Advances in prospect theory: Cumulative representation of uncertainty. Journal of Risk and Uncertainty 5, 297–323.

Additional tables

Table A: **Prospectus data**

Table A shows the number of prospectuses for each form type that mutual funds have to fill and whether the information is already at the fund level or the fund family level. 497, 497K, 485APOS, and 485BPOS filings are available in the EDGAR SEC database. In total, we retrieve fund names from 418,938 prospectuses, covering 20,973 funds (SeriesId's). Prospectus data is provided in the eXtensible Business Reporting Language (XBRL) format. The fund name in a prospectus can be identified due to the XBRL tree and key structure. We extract for all prospectuses the most recent name on a fund for each fund and month-year. Merging the data with the Morningstar database via the share class ticker symbol and restricting the sample to equity funds only results in 2,166 funds between 2010 and 2018.

Prospectus type	# of observations	Fund level
497	164,654	Yes
497K	168,409	Yes
485APOS	12,586	No
485BPOS	73,289	No
Total	418,938	

Table B: **Possible cases of inaccurate names**

Table B shows the eight possible cases that we have in our sample and the associated number of fund-quarter observations. A fund name either refers to the SMB dimension (*small* or *large*), to the HML dimension (*value* or *growth*), or to both dimensions. When a fund name refers to both dimensions, it can be inaccurate (or accurate) on only one of these two dimensions, or on both.

	Name	SMB dimension	HML dimension	N
(1)	SMB	Accurate		5,513
(2)	SMB	Inaccurate		135
(3)	HML		Accurate	10,752
(4)	HML		Inaccurate	1,961
(5)	SMB & HML	Accurate	Accurate	7,729
(6)	SMB & HML	Inaccurate	Accurate	116
(7)	SMB & HML	Accurate	Inaccurate	748
(8)	SMB & HML	Inaccurate	Inaccurate	48

Table C: **Control and treatment groups**

Table C shows the number of fund-quarter observations during which a name switches from being accurate to being inaccurate (columns *Treated*) or remains accurate (columns *Control*). The numbers are reported per type of name and per dimension on which a name is inaccurate. We obtain three different treatment and control groups, one for each dimension on which a fund name can be inaccurate. SMB refers fund-quarter observations associated to cases (1), (2), (5), and (6) in Table B. The fund names in the treatment group switch from being part of the accurate name group (1)+(5) to being part of the inaccurate name group (2)+(6), while fund names in the control group remain in group (1)+(5). HML refers to fund-quarter observations represented by cases (3), (4), (5), and (7). The fund names in the treatment group switch from being part of the accurate name group (3)+(5) to being part of the inaccurate name group (4)+(7), while the fund names in the control group remain in group (3)+(5). SMB & HML refers to fund-quarter observations of cases (5) and (8). The fund names in the treatment group switch from being part of the accurate name group (5) to being part of the inaccurate name group (8), while the fund names in the control group remain in group (5). Note that, when a name includes both dimensions and remains accurate on both, the associated fund-quarter observation is included in the SMB control group, in the HML control group, and in the SMB & HML control group.

	SMB		HML		SMB & HML	
	Control	Treated	Control	Treated	Control	Treated
Small	3876	100				
Small & Growth	1299	28	1299	224	1299	4
Small & Value	2496	42	2496	56	2496	16
Large	1524	14				
Large & Growth	1696	17	1696	43	1696	18
Large & Value	1661	15	1661	77	1661	3
Growth			5323	576		
Value			4448	379		

Table D: **Variable definitions**

Table D provides definitions of the variables used in this paper. MS indicates Morningstar, C refers to own calculation and SEC indicates data from the Securities and Exchange Commission.

Variable name	Description	Source
Name	Name history of the fund as extracted from all prospectus fillings in EDGAR	SEC
Series ID	Fund identifier as extracted from the prospectuses in EDGAR	SEC
Fund Flow (in %)	Computed as $(TNA_{f,t} - TNA_{f,t-1}(1 + RF_{f,t}))/(TNA_{f,t-1}(1 + RF_{f,t}))$, where $TNA_{f,t}$ corresponds to fund f's total net assets (TNA) in quarter t and $RT_{f,t}$ denotes fund f's return in quarter t. The variable is winsorized at the 1st and 99th percentile.	MS, C
Return (in %)	Quarterly log-return computed as the sum of monthly log-returns. Monthly returns are computed as the percentage return calculated as the change in monthly net asset value minus management fees and other regular costs.	MS, C
Rank (in %)	Rank of the fund return in comparison to funds having the same terms in their names. Fund return is the quarterly log-return.	MS, C
Idiosyncratic Risk (in %)	Standard deviation of the residuals from a model including the three Fama and French factor returns as well as the Carhart momentum factor. Factor returns are from Kenneth French's library.	MS, C
Fund Size	Logarithm of fund's quarterly total net assets in million USD, aggregated at the fund level.	MS, C
Company Size	Logarithm of company's quarterly total net assets in million USD, aggregated at the investment company level.	MS, C
Age (in Years)	Fund's age computed as the difference from quarter t to the inception date of the oldest share class.	MS, C
log(Age (in Years))	Logarithm of fund's age in months computed as the difference from quarter t to the inception date of the oldest share class.	MS, C

Variable definitions, Table D continued

Variable name	Description	Source
Expense Ratio (in %)	Fund's quarterly expense ratio expressed in percent.	MS
Turnover Ratio (in %)	Fund's quarterly turnover ratio expressed in percent.	MS
Inaccurate	Classification of fund's name as accurate (0) or inaccurate (1) based on the Fuzzy C-means clustering method in quarter t.	C
Cash Proportion	The percentage of the fund's net assets in cash.	MS
FI Proportion	The percentage of the fund's net assets in fixed income.	MS
Equity Proportion	The percentage of the fund's net assets in equity.	MS
Holdings	Quarterly number of holdings that are in the portfolio of a fund.	MS