

# Mutual funds styles, distinctiveness and financial performance: Insights from Europe

Sophie Béreau <sup>1,2</sup>, Jean-Yves Gnabo <sup>1</sup>, Henri Vanhomwegen <sup>1</sup>

<sup>1</sup> CeReFiM and naXys, Université de Namur

<sup>2</sup> CORE, Université catholique de Louvain

Very preliminary version, please do not quote or circulate

## Abstract

Are equity mutual funds following the same style over their lifespan? If not, how often do they change? How much a fund is distinct from its style-based peers? What are the drivers of mutual funds' distinctiveness? And ultimately does it pay to be different? Those are central questions for both academics and market practitioners that we aim to address in this paper. Drawing on the contribution of Sun et al. (2012), we propose to study the distinctiveness of European equity mutual funds over time, and its impact on financial performance. To this end, we use several market-based measures of fund-level distinctiveness on 4284 European equity mutual funds over the 1999-2016 period. Then we use regression analysis to unveil the determinants of equity mutual funds distinctiveness and its impact on the financial performance. We contribute to the literature in three ways: (i) we are the first to propose such analysis on European funds (ii) we use a fully dynamic approach to retrieve peers' groups and then compute the level of distinctiveness of each fund at each point in time, (iii) we apply the recent approach of Boudt and Ardia (2017) to formally test the difference of performance across peers'. Our results show marked changes in the level of commonality over time, distinctiveness among European mutual funds decreasing sharply before the crisis before going up and eventually reaching a new low in the recent months. We find strong evidence regarding the contribution of fund-level as well as cluster characteristics to strategy distinctiveness. Turning to financial performance, our results show a strong and robust, positive impact of strategy distinctiveness on financial performance.

J.E.L. classification: G11, G12, G23

Keywords: mutual funds, commonality, dynamic clustering

## 1. Introduction

Over the past fifteen years, the global asset management industry has more than doubled in size, reaching a total amount of \$71.4 trillion of assets under management (AuM) by the end of 2015 from \$29 trillion in 2002 (Boston Consulting Group 2016). What is even more remarkable is perhaps the fact that such development has been exacerbated in the aftermath of the financial crisis while other financial industries such as the banking industry, were experiencing substantial down-scale in their activity (McCord et al. 2015). According to the Boston Consulting Group (2016), the asset management industry raised by 38% between 2008 and 2014 (\$70.5 trillion) bouncing back from a short lasting contraction of nearly 20% in 2008<sup>1</sup>. Among all the types of funds constituting the industry, mutual funds account for the lion share with an estimated global size of \$35 trillion in 2015. While being historically barely considered as a significant source of risk for the financial system as opposed to banks and hedge funds for instance, mutual funds have recently called the attention of surveillance authorities expressing concerns about recent developments in their practices (International Monetary fund

---

<sup>1</sup> The total amount under management was \$52 trillion in 2007 and \$43 trillion in 2008.

2015; Financial Stability Board 2015). Two issues in particular have been discussed. The first one deals with the growing part of their credit intermediation activities that constitutes part of the so-called “shadow-banking”, making the industry more interconnected with the rest of the financial system. The second relates to the consequences of recent changes in the investment environment. Indeed, while mutual funds have been historically characterized by a low-risk profile, the low interest rate environment has encouraged more uniform and aggressive practices, reinforcing one-sided market and making them potentially riskier to the rest of the system. While a growing literature has used US data to document these new patterns in the industry along with their implications in terms of risk and performance (Becker et al. 2015), limited evidence has been provided for the European mutual funds so far although their total amount of assets under management reach \$9 trillion in 2014 (ICI 2015) that is half of the US mutual fund industry or 20% of total assets of the European banking industry for instance.

Against this background, our aim is to provide new insights on the mutual funds industry by documenting the distinctiveness of European mutual funds and its implication for mutual funds' performance. The notion of distinctiveness includes two dimensions: (i) a system-wide level dimension when each fund is compared to the whole universe of funds, (ii) a close-peers level when the comparison is restricted to what could be considered as the fund's main competitors. Like Sun et al. (2014) in the context of hedge funds, we are in this analysis mainly interested in the later type of distinctiveness. As discussed in Boudt and Ardia (2017), the peer category which plays a critical role in the assessment of funds performances is defined “in a sufficiently broad manner in order to create a level playing field where the peer funds still have ample degrees of freedom to distinguish themselves”. Our central question is therefore whether the strategy consisting in being distinct from the close peers is rewarded with higher financial performance. To address this question, we have constructed an original large-scale database containing monthly observations of 8517 (4284 after cleaning) European equity mutual funds over the 1999-2016 period. We compute different measures to detect the set of “close peers” to be considered by applying advanced clustering techniques in statistics and machine learning. Equipped with those measures, we use a regression analysis to document the determinants of equity mutual funds distinctiveness and its marginal contribution to financial performance. Our results provide clear evidence regarding the contribution of fund-level as well as cluster characteristics to strategy distinctiveness. Turning to financial performance, our empirical estimates confirm a strong and robust, positive impact of strategy distinctiveness on mutual funds financial performance.

In what follows, we briefly review the main strands of the literature on mutual funds related to our analysis before discussing our specific contribution with respect to existing studies.

The first strand of the literature has been stimulated amid the Financial Stability Board's proposal of systemic risk regulations for the asset management industry, among which mutual funds. It extends previous studies analyzing the potential contribution of hedge funds to systemic risk by Adrian (2007), Savona (2014) and Kaal and Krause (2016) to quote few. Hence, Roncalli and Weisang (2015) design a model to assess rigorously the candidates for the “systemically important” denomination of the FSB (i.e. the so-called “SIFIs”). To that end, they modify the factors proposed by the FSB to assess systemic risk allowing to flag more relevant funds. In his paper, Wan (2016) takes a broader perspective discussing the pros and cons of adopting such a labeling in the asset management industry. His main conclusion is that such labelling, while useful for the banking and insurance industry, is not well suited for

the asset management given the widely different business models of its constituent. Approaching the systemic importance's question of mutual funds from another angle, Béreau et al. (2015) investigate the temporal evolution of commonalities in total net asset returns across US equity mutual funds. Their idea is that increased homogeneity within the industry could create one-sided market and contribute to system-wide financial instability as measured by price distortion and tail events. Their findings tend to confirm the increasing risk posed by this industry to the rest of the system.

A second strand of the literature related to our study lies in fund performance and classification. In their study, Sun et al. (2012) investigate the link between unique strategies followed by hedge fund managers and the resulting performances. To do so, they provide a new measure of funds' distinctiveness called the "Strategy Distinctiveness Index" (SDI hereafter). The SDI measures how much each fund's returns differ from the average returns of their peers among specific groups characterizing their investment styles, those investment styles being previously defined by means of dynamic clustering techniques on returns data as done by Brown and Goetzmann (1997). A recent study by Vozlyublennaia and Wu (2016) extends the contribution of Sun et al. (2012) by introducing a two-level cluster analysis. First, hedge funds are separated into K-style clusters, then hierarchical clustering techniques (which produces among others a dendrogram representation) is applied inside each K-cluster. The main rationale behind the procedure is that the latest single fund to be linked to the dendrogram will be the most different or unique within its cluster. The results show that performance of that most unique fund is on average superior to peers' performance within the same cluster.

Drawing on existing researches, our study aims to contribute to the literature with respect to three main aspects.

First, while almost all studies have been performed on US asset management funds,<sup>2</sup> we use in this paper an original database on European equity mutual funds to shed light on an important segment of the asset management industry which has been so far under explored. The focus of the academic literature on the US case can be explained by several reasons. The size is a first argument as the total amount under management represents 50.1% of all AuM (Boston Consulting Group, 2016). A second argument lies in data availability. Indeed, data on fund characteristics are more easily accessible in the US as the various actors from this industry have to comply with the Security & Exchange Commission (SEC) strict transparency rules (Security and Exchange Commission, 2004). According to this rule, all registered open-end funds (except Money Market Funds) are required to disclose their complete portfolio holdings on a quarterly basis. Eventually, the legal environment is simpler in the US than for instance in other large markets such as Europe where multiple legal structures do apply. The case of Europe remains however of utmost importance for the financial industry as a whole and the mutual fund industry in particular. Indeed, the European market share represents 27.5% (of which 46% are mutual funds, ICI, 2015) of the global market which is substantial. In addition, if we take a temporal perspective, we can observe that the North-American market is

---

<sup>2</sup> Notable exceptions are Walter and Moritz Weber (2006) or Frey et al. (2014) which investigate the evidence of herding behaviors in the German mutual fund industry as well as Otten and Bams (2002) on European mutual funds performances

currently shrinking (-1% between 2014 and 2015), while its European counterpart is still exhibiting a steady growth (3% over the same period) (Boston Consulting Group, 2016).

Second, we contribute to the literature by making use of advanced machine learning techniques in line of Xu et al. (2014); Sarda-Espinosa (2016); or Charrad et al. (2014) to identify the set of relevant close peers over time. As developed in the pioneering contribution of Sun et al. (2012), there could be a mismatch between the style to which a fund is attached in the market and their actual investment strategy. To overcome this problem Sun et al. (2012) propose to retrieve the style by the mean of statistical methods by looking at co-movements in their net asset values. The expected outcome of this bottom-up approach is to provide a more accurate set of “close peers” to construct the measure of distinctiveness for each fund. An important dimension discussed in Sun et al. (2012) for the identification of peers is the dynamic nature of the fund’s strategy, with funds moving from one style to another over their lifespan. As often done in applied works, they propose to deal with this issue by using a rolling windows approach. Peers are identified over sub-samples that only include the most recent past observations. However, as widely known also in the econometric literature, rolling windows leaves researchers with no clear guidance to decide the length of the window and can be highly inaccurate to depict dynamic system especially in the presence of outliers. We do propose to contribute to the detection of close peers in the asset management industry by applying the Adaptive Forgetting Factor for Evolutionary Clustering Tracking (AFFECT) developed by Xu et al. (2014). This approach is purely dynamic and does not require to split the sample in ad-hoc windows.

Third, we are applying formal statistical tests to assess whether a fund out or underperform its peers. As discussed in Boudt and Ardia (2017) too often applied works directly consider excess returns stemming from the CAPM, Fama-French three factors or the Carhart model as an evidence of over/under performance with respect to other funds. Such a procedure can be misleading however as those returns are estimated quantities for which formal statistical tests are required to make pairwise comparison. We apply Boudt and Ardia (2017) recent procedure to recover a cleaner measure of performance for our European equity mutual funds and in turn test its link with the level of distinctiveness.

The remainder of the paper is the following, Section 2 details the data used in our study. Section 3 focuses on our empirical methodology by first exposing our various measures of funds commonalities relying on alternative clustering techniques, then our regressions devoted to documenting commonalities’ determinants and finally, commonalities’ impact on funds’ performance. In Section 4 we address some empirical limitations and provide a roadmap for future robustness checks and finally, Section 5 concludes.

## **2. Data**

Our data are retrieved from the Morningstar Direct Database with a focus on active equity mutual funds domiciled in Europe, traded in Euro and for which monthly returns are observed over the 1999-2016 period. In order to consolidate our database, we exclude funds with less than 10 monthly observations and with an average size of less than 10 million euros. Considering the multiple share classes offered by specific funds we select the one for which most observations are available, doing so, our database finally consists of 8517 different mutual funds’ portfolios. Drawing on Sun et al. (2012), we start our analysis by allocating the

whole set of funds into sub-groups based on styles, keeping only categories of funds containing at least a hundred of them, leading to a final sample of 4284 portfolios. To that end, we use the categories proposed by Morningstar. In a second step, we use a more sophisticated identification strategy consisting in applying statistical and machine learning techniques on mutual funds' Total Returns (TR) to extract dynamically clusters of funds from the time series data (see Section 3.1 for more details on our empirical methodology).

Based on the monthly total net assets (TNA) and returns (R) data at the portfolio level, we compute mutual fund flows (F) as in Coval and Stafford (2007):

$$F_{i,t} = TNA_{i,t} - TNA_{i,t-1} * (1 + R_{i,t})$$

Our current focus is on active equity mutual fund portfolios domiciled in Europe trading in euro currency. The decision to focus on these funds is threefold. First, this market has been poorly documented by academic research. Second, we are able to study these funds over time since the end 1990s. Third, we capture a reasonable market share of the trades in European equity market as detailed below, capturing mentioned. For this group of mutual funds, a sufficient sample size is obtained for each month over the 1999-2016 period. In total, there are 5930 portfolios in our sample corresponding to categories of funds including at least a hundred of them. At end-2016 the total TNA of these active equity mutual funds domiciled in the euro area and trading in euro currency was about 1.191 trillion euros. The whole mutual fund industry registered in Europe amounted to 2.462 trillion euros at end-2016<sup>3</sup>, so our sample captures a bit more than 48% of all the euro-area domiciled equity funds trading in euros.

To investigate the drivers of strategy distinctiveness of equity mutual funds, several fund-level characteristics variables, cluster-level variables as well as global financial indicators are collected: (i) the size of the funds as measured by its total net assets value (TNA) discussed above, past volatility of net asset returns, net flows, the age of the funds, its past excess returns, bull beta and bear beta indicators, (ii) the average return and the cross-sectional dispersion of total net asset returns of the funds included in the same cluster along with the overall total net asset values of the cluster, (iii) and eventually the VStoxx.

Next to the mutual fund data and the variables that serve as possible determinants of commonality, the analysis also involves return data of equity indices to estimate the risk adjusted performance of equity mutual funds. The regression used to explain the performance are the following: (i) fund size, fund volatility, excess return and the fund age as done in the SDI equation along with (ii) the fund flow, the monthly price-to-book (PB) and price earnings (PE) ratios, the monthly average return of the fund, the average drawdown and eventually the longest number of consecutive positive and negative results (in periods).

### 3. Empirical exercise

---

<sup>3</sup> The detailed figures of aggregated balance sheet of euro area investment funds broken down by investment policies provided by the statistical data warehouse of the ECB are available here:

<http://sdw.ecb.europa.eu/reports.do?node=1000003520>

In this section, we detail our empirical approach to first assess commonalities among our sample of European mutual funds (Section 3.1), then estimate their potential drivers (Section 3.2) and finally quantify the impact of time varying commonalities on mutual funds' performance (Section 3.3).

### 3.1. Measuring mutual funds distinctiveness

In order to capture the degree of distinctiveness among European equity mutual funds' investments at both the system wide and the individual levels, we follow the approach proposed by Sun et al. (2012) consisting in using the "distance" between funds' net asset returns. The underlying motivation for using such a proxy is simply that we do expect closer net asset returns for funds having similar portfolio and then similar strategy. An alternative would have been to rely directly on portfolio holdings. However, net asset returns are available at a higher frequency providing then more flexibility to document the evolution of commonality over time. Second, working with portfolio holdings generally involves a high price, as the data collection is long and heavy.

As noted above, our main measure is constructed by following the approach developed by Sun et al. (2014) to compute the strategy distinctiveness index (SDI). In their contribution, this index is aimed to assess how distinct and unique a hedge fund strategy is compared to its peers. Formally, the SDI measure is calculated for each fund  $i$  as follows:

$$SDI_{i,t} = 1 - corr(r_{i,t}, \mu_{I,t})$$

$$= 1 - \frac{\sum_{t=1}^{24} (r_{it} - \bar{r}_i)(\mu_{It} - \bar{\mu}_I)}{\sqrt{\sum_{t=1}^{24} (r_{it} - \bar{r}_i)^2 \sum_{t=1}^{24} (\mu_{It} - \bar{\mu}_I)^2}}, \quad \mu_{I,t} = \frac{\sum_{(i \in I)} r_{it}}{count(i \in I)}$$

This metric corresponds thus to 1 minus the correlation between the hedge fund's returns ( $r_{it}$ ) and the average returns of all funds belonging to the same cluster or style indexed by  $I$ . The higher the SDI, the more distinct the hedge fund's investment strategy with respect to its cluster.

One critical aspect of the methodology is to identify the fund style. Indeed, a fund is not deemed as distinct with respect to the whole universe of funds but is compared to its peers, its peers being those adopting the same style. To compute the measure, we therefore need to identify in a first step the different styles at stake in the sample. To that end, we follow two strategies. First, we rely, as Sun et al. (2012), on style categories as reported by the private database Morningstar. Second, we infer the styles from the data by applying several statistical and machine learning techniques. We develop these approaches in the next subsection.

The allocation of mutual funds into the accurate similarity-based category is a not an easy task. Yet, such information can be extremely valuable either for academic research or for professional investors (e.g. help identify the real benchmark the fund should be compared to). In order to group our sets of funds in a consistent manner, we apply up-to-date clustering algorithms on funds' net asset returns. The purpose of the clustering approach is to measure similarities between observations and to gather them in clusters by applying an optimization

algorithm. Two methods appear particularly well suited for our purpose: partitional algorithms and hierarchical algorithms.

The main approach falling into hierarchical algorithms is the so-called “K-mean” algorithm (Hartigan and Hartigan, 1975). It aims at grouping observation into ‘k’ clusters specified ex-ante, where each observation belong to the closest cluster mean. The process is traditionally iterative. First, k observations of the dataset are assigned to be the initial cluster means. Second, observations are gathered according to their nearest mean which is computed as the smallest Euclidean distance. It then proceeds by iterating between step-one and step-two until results converge to a final segmentation of the data.

The main partitional algorithm is the so-called “agglomerative” hierarchical algorithm. It is a bottom-up iterative approach that first links the two closest observations together. Then, it iterates by linking the two closest elements together (i.e observation-observation, observation-cluster or cluster-cluster) until all individual observations belong to one cluster, creating a ‘tree’ called a dendrogram. The number of clusters is thus specified ex-post after interpretation of the dendrogram. To compute the distance, a symmetric square matrix of pairwise similarities between all observations is needed (dissimilarity matrix). For more information on the distance and linkage measures, see Gan et.al (2007).

An important critic that should be raised regarding the two previous algorithms lies in their static nature as they have been designed to deal with non-dynamic systems. In our case, it is highly likely that the funds change their strategy over time following a certain style in a period of their life before moving to another one, adjusting to varying market conditions and perceived profitability. A way to deal with the time-varying nature of styles and then clusters, is to use a rolling window approach as done in Sun et al. (2012). In this framework, cluster algorithms are applied to successive sub-sample windows. As discussed in the statistical literature however such an approach can provide a misleading picture of the system over time. The outcome for instance is highly sensitive to entry and exit of outliers from the window. Over the recent years, alternative approaches have been proposed by the literature to deal with this issue in the context of cluster identification. Here, we consider as an alternative to rolling-window estimations, an algorithm rooted in machine learning, called “AFFECT algorithm” (Xu et al, 2014). This algorithm has been specifically designed to deal with dynamic clusters. The AFFECT algorithm is iterative and computes a matrix of dissimilarity in  $t=1$  to which it applies a static agglomerative hierarchical clustering method. In the following steps, the dissimilarity matrices obtained will be revised to take into account past information thanks to the adaptive forgetting factor (Xu et al, 2014), this allows to use the static partitional algorithm yet retrieve dynamic results.

Equipped with those measures, we can compute alternative SDI measures in which the cluster of reference for each fund (i.e. group of peers following the same style) is potentially different, our preferred measures being the one resulting from the application of the AFFECT procedure. Graphs A.1-A.5 presented in Appendix feature the evolution over time of the computed SDI index based on alternative methods previously defined, on raw vs. normalized data. Although similarities are more striking for SDI index obtained from the first two methods i.e. “K-means” and hierarchical methods, since they lead to very close computations, whereas the time-varying index stemming from AFFECT dynamic clustering method presents more singularities, common patterns can be observed for all three, which include a significant increase in

commonalities before the crisis, followed by a sharp decrease over the end-2005 early-2011 period and a second rise at the end of our sample.

### **3.2. Determinants of commonalities**

Equipped with the different measures of mutual funds' distinctiveness (resp. commonality) previously discussed, we then use a regression analysis setting to shed lights on their main determinants that is to say to identify which factors lead a mutual fund to follow a strategy potentially different from its peers. The set of variables considered for this analysis can be broken down into three categories: (i) fund-level variables, (ii) cluster-level variables as well as (iii) global variables. Following Sun et al (2012), we apply a multivariate panel regression setting to monthly data. A within transformation is used to control for funds fixed effect. In addition, to mitigate potential endogeneity problem, the determinants are taken with one-period lag.

In what follows, we list the whole set of variables included into our models. As noted above, the first set pertains to fund-level characteristics. Hence, we include the asset under management (Fundsize), the fund net asset returns volatility (VolPast2Y), the age of the funds expressed in years (age), the net flow of funds expressed in percentage of the fund size (FlowPast2Y), their excess returns and eventually their sensitivity to positive/negative changes in their benchmark's return (Bull Beta/Bear Beta). The second set of variables concerns the cluster the funds belong to. We include the size of the cluster (clustersize) estimated by the ratio between the number of peers within the same cluster and the total number of funds at the same period (in percentage), the mean return of the cluster (clustermean) and its dispersion as measured by the standard deviation across funds returns (clusterstd). Last but not least, we take into account systematic

Table 1  
Determinants of the SDI (1999-2016)

SDI	(I)	(II)
	K-mean	K-centroids
Fund Size(ln)	-0.003***	-0.003***
Fund Volatility	-0.005***	-0.006***
Flow	.009***	.007***
Age (Year)	-0.001***	-0.001***
Cluster mean return	-0.001	.001
Cluster heterogeneity	.552***	.357***
Cluster size	.001	.001
Global volatility	-0.001***	-0.001***
R2	67.05	63.12
F-stat	190***	191***
#Funds	4,267	4,267

Table 1 reports the fixed effect panel regression of the SDI on cluster characteristics, a macro variable and on fund characteristics using monthly data. The equation is of the following form:

$$SDI_{i,t} = \alpha_i + \beta' \text{Fund characteristics}_{i,t-1} + \delta' \text{Cluster characteristics}_{i,t} + \eta \text{VStoxx}_t + \varepsilon_{i,t}$$

SDI, the volatility of return (VolPast2y), the estimated net flow of the funds (FlowPast2Y), are measured over a window of 24 months at the exception of SDI, lagged by one month. The funds size, excess return and age are derived monthly and lagged by a month. The cluster characteristics consist of the average return of the clusters (clustermean), the standard deviation of the clusters (clusterstd) and % of the total number of funds inside each clusters (clustersize). Columns I&II report the results of the K-mean cluster algorithm on normalized time-series and of Hierarchical clustering on normalized time-series respectively. Results for non-normalized series are qualitatively similar and remain available upon request. Results are adjusted for fund clustering effect and time fixed effects through VStoxx a European volatility index based on the Eurostoxx50 market. \*\*\*, \*\*, \* stands for statistical significance at 1%, 5% and 10% levels respectively.

factors in distinctiveness by including a global variable, namely the VStoxx which is a standard measure of uncertainty in financial markets.

Table 1 reports the results for the different measures of SDI. In the first column, the k-mean algorithm on normalized data is used for detecting the clusters and then to compute the SDI. In column 2, the hierarchical algorithm is applied as an alternative to detect the clusters on normalized returns. The first notable feature comes from the F-stat as we can strongly reject the null hypothesis of absence of significance for the whole set of regressors. Looking at individual variables. Other interesting features emerge. Hence, all funds-level variables at the exception of excess returns influence significantly the SDI. This result is very strong as the null hypothesis can be rejected in all cases at 1% and fully consistent across the four models. The sign of the coefficient is negative for the variable depicting the size of the fund, its age or the volatility of its net asset returns and sensitivity to benchmark. These results are consistent with Sun et al (2012) findings on Hedge funds. The volatility and sensitivity to benchmark results are quite straightforward: if a fund has been facing an increase in uncertainty (high volatility in its returns) it might decide to lower risk exposure and then to scale down its differentiation strategy.

The interpretation for the remaining variables, fund age and size is less straightforward. As it stands, the negative sign of the coefficient means that large and old funds are less prone to follow a distinctive strategy. Several mechanisms could be at play to explain this result. For instance, increased pressure from clients with the size and the age could limit the ability of the fund to follow an original strategy since as stated by Sun et al. (2012) obtaining similar results for hedge funds, SDI could also be viewed as reflecting a talent for innovation, which makes those results consistent with the idea that young funds are more likely to pursue innovative ideas while managers of small funds being more flexible can more readily incorporate innovations into their current practice. It is important to note however that these conjectures about the channels through which size and age affect the SDI cannot be formally tested within our current framework. Finally, funds that are more active are able to differentiate from others as suggested by the fund flow coefficient, which is consistent with the idea that differentiation is not effortless but rather due to specific investment strategies. Turning to the variables characterizing the cluster, we can also notice strong significance for one variable out of three. Hence, consistently with what was expected the dispersion within the cluster does have a positive effect on SDI. The conclusions are more mitigated for the cluster size as the variable is significant only in one case.

Eventually, our global variable VStoxx significantly influence the level of distinctiveness. As the sign is negative, it means that increased uncertainty leads mutual funds to adopt similar strategies all things being equal.

### **3.3. Quantifying the impact of commonalities on funds' performance (Very preliminary)**

In a second step, our aim is to document the link between the level of distinctiveness and the performance of mutual funds. By doing so, we want to quantify whether original strategies are in general rewarded with higher returns. The SDI now enters the model as a regressor. Our dependent variable is a measure of mutual fund performance. To that end, we take a measure of Jensen's alphas which is commonly used in the literature to determine the abnormal returns on a portfolio of securities over the theoretical expected one based on standard modelling

(CAPM or alternative multi-factor models). To allow time variation of performance, our alternative measures are computed as monthly realized alpha from daily returns that is at each monthly period our alpha is re-estimated using within month daily observations. Using this raw information to construct our dependent variable could be misleading as two estimated alpha could be different while not statistically different from one another, the divergence of values being due to data sampling. To obtain a more reliable measure of outperformance (resp. underperformance) of an equity mutual fund in its peers universe, we use the testing procedure proposed by Boudt and Ardia (2017).

Equipped with these variables, a panel fixed effect model is adopted with a set of control variables in addition to the SDI.

As done in Sun et al. (2012), we include among the controls, the characteristics of mutual funds appearing in the SDI equation: fund size, the fund flow and the fund age. In addition, a set of variable pertaining the cluster to which the fund is attached is included: the cluster mean return, the cluster heterogeneity as the standard deviation of returns of funds included in the cluster, the cluster size.

Table 2 reports our results from the estimation of multivariate panel regressions for raw measure of alpha (col. 1) and the outperformance measure proposed by Boudt and Ardia (2017) (col. 2). We estimate in each case a static and a dynamic version, the later including the first lag of the dependent variable among the regressors.

The main important result is that the variable SDI is positive and significant in the different models which support the hypothesis that strategy differentiation across funds' do impact positively their performance. In addition, we can observe from auxiliary variables that the funds' performance turns out to be negatively correlated with the size, flow and age.

Table 2  
Panel regression of Mutual Fund performance on SDI (1999-2016)

	<b>ALPHA CAPM</b>			
	(I)	(II)	(III)	(IV)
	<b>Alpha</b>		<b>Outperformance</b>	
Alpha (Lag1)		.001		-.006
SDI	.008**	.008**	.030***	.031***
Fundsize	-.002***	-.002***	-.003***	-.003***
ageyear	-.001**	-.001**	-.001***	-.001***
Flow	-.001	-.001	.003	.003
Vstoxx	-.001***	-.001***	-.001***	.001***
Cluster controls	Yes	Yes	Yes	Yes
F-Stat	11.5***	10.6***	46.3***	43.4***
#Funds	4284	4284	4243	4243

Table 2 reports the fixed effect panel regression of funds' performance captured by Jensen's alphas on SDI and a set of controls including cluster characteristics, a macro variable as well as fund characteristics using monthly data. The equation is of the following form:

$$\alpha_{i,t} = c_i + \beta_0 \text{SDI}_{i,t} + \beta' \text{Fund characteristics}_{i,t-1} + \delta' \text{Cluster characteristics}_{i,t} + \eta \text{VStoxx}_t + \varepsilon_{i,t}$$

As previously, SDI, the estimated net flow of the funds (FlowPast2Y), lagged by one month. The funds size, excess return and age are derived monthly and lagged by a month. Additional variables from previous SDI regression include fund size, fund volatility, excess return and the fund age. The cluster characteristics consist of the average return of the clusters (clustermean), the standard deviation of the clusters (clusterstd) and % of the total number of funds inside each clusters (clustersize) and are introduced in levels as well as in interaction with SDI measure. Columns I&II, V&VI report the results of the performance being measure as the estimated Jensen alpha or the outperformance measure proposed by Boudt and Ardia (2017) respectively. Results are adjusted for fund clustering effect and time fixed effects through VStoxx a European volatility index based on the Eurostoxx50 market. \*\*\*, \*\*, \* stands for statistical significance at 1%, 5% and 10% levels respectively.

Several robustness tests have been performed by addition other control variables, changing the sample or using alternative models for computing the Jensen's alpha. The main conclusion remain unchanged to these modifications of the estimation procedure.

## 5. Conclusion and discussion

This paper studies the impact of European equity mutual funds distinctiveness with respect to close peers on performance. To do so, we construct a database of 4284 European equation

mutual funds operating during the 1999-2016 period. We pay a particular attention to the computation of adequate metrics, the so-called « strategy distinctiveness index » applying up-to-date techniques from statistics and machine learning to identify the “close peers” to be compared to for each fund. In addition, we use the recent statistical procedure developed by Boudt and Ardia (2017) to better measure the relative performance of funds. Our preliminary results show marked changes in the level of commonality over time, distinctiveness among European mutual funds decreasing sharply before the crisis before going up and eventually reaching a new low in the recent months. We also find strong evidence regarding the contribution of fund-level as well as cluster characteristics to strategy distinctiveness. Turning to financial performance, our empirical estimates confirm the existence of a strong, robust, and positive effect of strategy distinctiveness on financial performance.

## References

Adrian, T. (2007). Measuring risk in the hedge fund sector. SSRN.

Arellano, M. and S. Bond (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations, *The Review of Economic Studies*, 58(2), 277–297.

Boudt, K and Ardia, D. (2017) The Peer Performance Ratios of Hedge Funds. *Journal of Banking and Finance* (forthcoming)

Boston Consulting Group (2016). Global asset management 2016: Doubling down on data.

Becker, B. and Ivashina, V. (2015). Reaching for yield in the bond market. *The Journal of Finance*, 70(5):1863-1902.

Béreau, S., Casteleyn, A., Gnabo, J.-Y., and Zwinkels, R. (2015). Commonality among equity mutual funds: impact on financial stability and price discovery. Working Paper UNamur.

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

Charrad, M., Ghazzali, N., Boiteau, V., Niknafs, A., and Charrad, M. M. (2014). Package nbclust. *J. Stat. Soft*, 61:1-36.

Financial Stability Board (2015). Assessment methodologies for identifying non- bank non-insurer global systemically important financial institutions.

Frey, S., Herbst, P., and Walter, A. (2014). Measuring mutual fund herding—a structural approach. *Journal of International Financial Markets, Institutions and Money*, 32:219-239.

ICI (2015). 2015 investment company factbook : A review of trends and activities in the U.S. investment company industry.

International Monetary Fund (2015). Global financial stability report: Fostering Stability in a Low-Growth, Low-Rate Era.

Kaal, W. A. and Krause, T. A. (2016). Hedge funds and systemic risk. Handbook on Hedge Funds, Oxford University Press. Available at SSRN: <https://ssrn.com/abstract=2748096>

McCord, R., Prescott, E. S., Sablik, T. (2015). Explaining the decline in the number of banks since the great recession. Richmond Fed Economic Brief, (March):1-5.

Nickell, S.J., (1981). Biases in Dynamic Models with Fixed Effects, *Econometrica*, vol. 49(6), 1417-1426.

Otten, R. and Bams, D. (2002). European mutual fund performance. *European Financial Management*, 8(1):75-101.

Roncalli, T. and Weisang, G. (2015). Asset management and systemic risk. In Paris December 2015 Finance Meeting EUROFIDAI-AFFI.

Sarda-Espinosa, A. (2016). Comparing time-series clustering algorithms in r using the dtwclust package.

Savona, R. (2014). Hedge fund systemic risk signals. *European Journal of Operational Research*, 236(1):282-291.

Security and Exchange Commission (2004). Shareholder reports and quarterly portfolio disclosure of registered management investment companies : Rin 3235-ag64.

Sun, Z., Wang, A., and Zheng, L. (2012). The road less traveled: Strategy distinctiveness and hedge fund performance. *Review of Financial Studies*, 25(1):96-143.

Vozlyublennaia, N. and Wu, Y. (2016). Mutual funds apart from the crowd. mimeo

Walter, A. and Moritz Weber, F. (2006). Herding in the German mutual fund industry. *European Financial Management*, 12(3):375-406.

Wan, J. S. (2016). Systemically important asset managers: Perspectives on dodd-frank' s systemic designation mechanism. *Columbia Law Review*, 116(3):805-841.

Xu, K. S., Kliger, M., and Hero III, A. O. (2014). Adaptive evolutionary clustering. *Data Mining and Knowledge Discovery*, 28(2):304-336.