

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/318921570>

The Impact of Benchmark Choice on US Mutual Fund Benchmark-Adjusted Performance and Ranking

Article in SSRN Electronic Journal · August 2017

DOI: 10.2139/ssrn.3014010

CITATIONS

2

READS

691

3 authors:



Irina Bezhentseva Mateus

Aalborg University

22 PUBLICATIONS 77 CITATIONS

[SEE PROFILE](#)



Cesario Mateus

Aalborg University Business School

88 PUBLICATIONS 373 CITATIONS

[SEE PROFILE](#)



Natasa Todorovic

City, University of London

35 PUBLICATIONS 122 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Value-Added Tax: Portuguese Experience and Anticipated Impact on the Greek Tourism Industry [View project](#)



A guide to survival of momentum in UK style portfolios [View project](#)

The impact of benchmark choice on US mutual fund benchmark-adjusted performance and ranking

Irina B. Mateus^a,

Cesario Mateus^a,

Natasa Todorovic^{b*},

^aUniversity of Greenwich, London, UK

^bThe Centre for Asset Management Research,

Cass Business School, City, University of London, UK

Abstract

This study re-visits the question of benchmark mismatch among 1281 US equity mutual funds and its impact on benchmark-adjusted fund performance and ranking. All funds report S&P500 index as a prospectus benchmark, yet 2/3 of those are placed in the Morningstar category with risk and objectives different to those of the S&P500 index. We identify ‘true’ benchmarks for those mismatched funds and find that their S&P adjusted alphas are higher than ‘true’ benchmark adjusted alphas in 61.2% of the cases. In terms of fund quartile rankings, 30% of winner funds lose that status when the prospectus benchmark is substituted with a more suited one. In the remaining performance quartiles there is no clear advantage of using S&P 500 as a prospectus benchmark. The prospectus benchmark therefore can mislead investors about fund’s relative performance. This leads us to conclude that any reference to performance in a fund’s prospectus should be treated with caution.

Keywords: Prospectus benchmark selection, Mutual fund benchmark mismatch, Benchmark-adjusted alphas, Performance ranking

JEL: G11, G12, G23

*Corresponding author. Tel.: +44 207 040 0120. E-mail address: n.todorovic@city.ac.uk (N. Todorovic).

1. Introduction

SEC regulations require mutual fund companies to disclose their performance relative to a passive benchmark, an index often referred to as their prospectus benchmark. Over a third of US investors rely on information in the fund prospectus when purchasing a mutual fund¹. Prospectus benchmark defines an investment direction and a risk tolerance, and should reflect the strategic role of the individual asset classes in the fund. However, Cremers and Petajisto (2009) provide evidence that mutual funds typically have a high proportion of holdings that differ from those of fund's (theoretically adequate) benchmark index. Sensoy (2009) affirms that funds frequently differ significantly from their benchmarks and shows that value funds are more likely to have self-designated benchmarks that are mismatched on value/growth, while small-cap funds tend to have prospectus benchmarks mismatched on size.

It should not come as a surprise then that some prospectus benchmark choices may be misleading, as there are no precise requirements on the selection of funds' best suited benchmark. Therefore, the choice of fund benchmark may be biased and may indicate principal-agent problems. As a consequence, for instance, a fund reporting a large cap index as their prospectus benchmark may have significant proportion of their assets invested in smaller size stocks. Considering investors' close scrutiny of fund performance it is vital to examine the extent of benchmark misclassification in US active fund management. Moreover, considering the development of recent literature on mutual fund performance, it is crucial to account for non-zero benchmark alphas, which significantly bias outcomes of fund performance (see for instance Chinthalapati et al., 2017). A recent study by Cremers, Petajisto and Zitzewitz (2012) shows that standard benchmark models produce economically and statistically significant non-zero alphas for passive benchmark indices, including a widely used US passive benchmark - the S&P 500. Negative and statistically significant alpha for the Russell 2000 Growth index was documented by Chan, Dimmock, and Lakonishok (2009); significant non-zero alphas are also discussed in Costa and Jakob (2006). The non-zero alphas of passive benchmarks are not solely a US phenomenon. Recently, persistent negative alphas are documented in the FTSE 100 Index in the UK (Mateus et al., 2016).

¹ Investment Company Institute, Understanding Investor Preferences for Mutual Fund Information, Summary of Research Findings ("Understanding Investor Preferences"), 2006, available at https://www.ici.org/pdf/rpt_06_inv_prefs_full.pdf

Based on the above, this paper aims to examine to which extent the benchmark choice of US long only equity funds changes inferences on fund performance, once the benchmark alphas are accounted for in fund performance evaluation. In particular, we assess whether inadequate prospectus benchmark selection may lead to over estimation of fund performance and whether it could be a subject of gaming. Further, we investigate whether benchmark choice affects fund performance in relative terms (relative to peers) and, therefore, changes the ranking position of the winning and losing funds, in particular. Hence, as our main contribution, we add to the literature on US mutual fund benchmark mismatch by 1) investigating the impact of the choice of benchmark on fund performance and performance rankings and 2) providing performance assessment free of biases caused by alphas embedded in the benchmark index and not accounted for in the standard pricing models. To account for these non-zero benchmark alphas, we apply Angelidis, Giamouridis and Tessaromatis (2013) methodology that allows for the alpha in the benchmark index to be included in a standard factor model, such as Carhart (1997). This approach adjusts alpha of a fund by that of the benchmark.

In the aspect of previous literature relevant to analysis, Sensoy (2009) provides evidence that funds frequently differ from their benchmarks in terms of their risk characteristics and composition for strategic reasons. Substantial exposures to size and value/growth factors in returns that are not captured by their benchmarks were also discussed in Elton, Gruber, and Blake (2003). The study of DiBartolomeo and Witkowski (1997) examine monthly returns for 748 load and no-load open-end funds and show that return patterns of 40 percent of funds analysed deviate from the benchmark declared in the prospectus with 9 percent of funds being seriously misclassified, two or more risk tiers away from their declared categories. Similarly, Kim, Shukla and Tomas (2000) assess how well mutual funds' stated objectives conform to their attributes-based objectives and revealed that the stated objectives of more than half the 1043 funds analysed differ from their attributes-based objectives, and over one third of the funds are severely misclassified. The study also confirms upward and downward risk shifts. Bams, Otten, and Ramezanifar (2016) analyse a sample of 1,866 US equity funds over the 2003-2015 period and found that 14% of funds are significantly misclassified based on long term style analysis. Huang et al. (2011) show that mutual funds change their total risk exposure substantially over time. Authors claim that it might be done for strategical reasons: in order to increase the expected money inflows to the funds or to manipulate their performance numbers. Similarly, Kacperczyk, Sialm, and Zheng (2008) measure the return gap, the difference

between the reported fund return and the return on a portfolio that invests in the previously disclosed fund holdings, and document that despite disclosure requirements, mutual fund investors do not observe all actions of fund managers. Portfolio performance manipulation and deviation from benchmarks was also discussed in Goetzmann et al. (2007), Jiang et al. (2014), Fung and Hsieh (2002).

This paper contributes to the mutual fund performance measurement literature. In addition it adds to the literature on mutual fund benchmark misclassification and extends the work of Chan, Dimmock, and Lakonishok (2009), which demonstrates that judgments about the magnitude of performance are sensitive to the benchmarking methodology. To the best of our knowledge, this is the first study that analyses the impact of benchmark choice on US equity fund performance and ranking while accounting for the non-zero alpha bias in those passive benchmarks. We use the net monthly returns of 1281 actively managed US equity mutual funds from January 1992 to February 2016. All funds in the sample report S&P500 as their primary prospectus benchmark in the Morningstar database. Our funds belong to 22 distinct Morningstar global categories: e.g. US Small Cap, US Large Cap Value, Energy Sector Equity, Global Equity etc. Investigation of commonly used benchmarks amongst funds in different categories in the Morningstar database, shows us that the primary prospectus benchmark that all our funds use, the S&P 500 Index, is most suitable for the funds in the Large Cap Blend Morningstar category. However, around 2/3 of the funds in our sample are not in that category, yet they declare S&P 500 as their passive benchmark. Our analysis of prospectus benchmark fit shows that the funds' rationale for selecting a particular passive index as prospectus benchmark is not clear, as the index does not correspond to funds composition or investment objectives in large proportion of our sample. For each of the Morningstar global categories, we identify a more appropriate benchmark than the S&P 500, which we refer to as 'true' benchmark in this paper. We find that 'true' benchmarks are a better fit for our funds than their prospectus benchmark, the S&P 500 index, having on average around 10% higher R-squared in the full sample period and each of the sub-periods. This makes an inference that even adjusted for benchmark alphas fund performance may be significantly biased if fund performance estimated versus unsuitable/prospectus benchmark is used by investors as a performance target.

To measure fund performance and rank the funds we apply Angelidis et al. (2013) methodology (AGT hereafter) that adjusts fund's alpha for benchmark's alpha, hence isolating manager's

skill above that common to the benchmark. We find that 61.2% of the mutual fund AGT alphas are higher when S&P500 is used as a benchmark². Further, in 15 out of 22 rolling periods of 36 months each, pairing the performance with S&P500 is beneficial to the funds and leads to overestimated performance. Thus, on average, prospectus benchmark amplifies fund performance by 23 basis points versus the performance adjusted with a ‘true’ benchmark. Nevertheless, there is still the remaining 30 percent of periods when performance is better when the ‘true’ benchmark alpha is used as the target in AGT model.

Analysis of fund quartile rankings shows that, on average, around 30% of winners leave the top quartile of funds when the benchmark is changed from the self-designated benchmark, S&P 500, to the ‘true’ benchmark in AGT benchmark-adjusted alpha estimation. On the opposite end of spectrum, nearly 30% of losers move up the quartiles when the ‘true’ benchmark is used. This shows that our results support the notion from Sensoy (2009) that the funds that appear at the top end of the spectrum may choose their prospectus benchmarks strategically. However, inappropriately chosen prospectus benchmark actually harms the funds that are at the bottom of the ranks. Given this, we conclude that the choice of the appropriate benchmark is critically important, as the wrong benchmark does not only bias performance assessment but can also lead to false conclusions when performance of funds relative to peers is assessed. Hence, this paper is of significant importance to individual investors, institutional investors and professional financial advisors interested in performance evaluation and fund rankings. Moreover, it has implications for financial regulators and policy makers with respect to fund information disclosure requirements and transparency in benchmark selection.

This paper proceeds as follows: Section 2 describes the data. Section 3 provides preliminary analysis where we test the existence of benchmarks alphas and check which benchmarks has a better explanatory power to the fund investment style. Section 4 presents Methodology. Section 5 analyses funds’ AGT-adjusted alpha performance and provides results. Section 6 delivers outcomes on biases in quartile ranking. Section 7 concludes.

2. Data

² The results presented are obtained with the use of the Carhart model in AGT augmentation. The outcomes obtained with Fama-French three and five factor models are qualitatively the same and available upon request.

The data set is comprised of 1,281 long-only active US equity mutual funds from January 1992 to February 2016. The net monthly returns of mutual funds are from Morningstar, inclusive of dividends. All funds have minimum requirement of 36 months of returns to be included in the sample. There is no survivorship bias in the sample. All funds in the sample declare S&P500 as their prospectus benchmark. However, due to the fact that in some cases investment objective stated in a fund's prospectus may not reflect how the fund actually invests (Kacperczyk, Sialm, and Zheng, 2008, Sensoy, 2009, Huang et al. 2011) Morningstar offers own proprietary data where each fund is assigned to a Morningstar Global Category based on the underlying holdings and fund's portfolio statistics. The list of Global Categories Morningstar has assigned our funds to, the number of funds per each category, the most relevant passive benchmark³ for each category and the number of monthly observations per category are presented in Table 1. The returns data for all benchmarks is inclusive of dividends.

----Table 1---

Only 36% of our sample (460 funds) fall in the Large Cap Blend Morningstar Global category where the S&P 500 would be deemed as the most appropriate passive benchmark. It means that performance analysis where the fund performance is measured against a prospectus benchmark can be biased and can provide inaccurate inferences about manager's skill. Further, 32% of our funds belong to the Large Cap Value and Large Cap Growth Global category where most commonly used benchmark for funds are Russel 1000 value and Russell 1000 Growth index respectively. Midcap Global Category is also highly presented (112 out of 1,281 funds, 8.7%). Its better fit would be Russell Midcap Index rather than S&P 500. Some of our funds are in the Small Cap Morningstar Global category (40 of 1,281), best represented by a Russell 2000 index. Overall, these aforementioned five categories (whose appropriate benchmarks should be S&P 500, Russel 1000 Value, Russel 1000 Growth, Russel Midcap and Russell 2000 index) account 80 percent of our sample. All other Morningstar Global categories in our sample are sector specific or country/region specific and call for sector or regional benchmarks. These specialist funds account for the remaining 20% of our sample, Hence, significant proportion (64%) of our funds selects and reports a benchmark inappropriate for their category of funds. This is important from two perspectives: 1) measuring fund performance relative to the

³ We review the indices that all available US equity funds in a given Global category benchmark against and select the most common benchmark, ensuring that its characteristics correspond to the category it represents (US Large cap value is best represented by Russell 1000 value Index etc.)

benchmark and 2) measuring fund performance relative to other similar funds. To this end, it is important to investigate fund's relative rankings within the same category and assess whether the funds that are the top performers according to prospectus benchmark (S&P500) change their relative ranking position after their performance is calculated with a benchmark that better reflects the risk characteristics of their Morningstar Global Category. Section 3.1 provides further discussion on suitability of the funds' self-declared benchmarks.

We split our analysis in 22 rolling overlapping sub-periods, each being 36 months of length. Given that the minimum data requirement for each fund is 36 months, within each rolling period we require that a fund has no less than 30 months of continuous returns. Table 2 reports the number of funds and monthly observations for each of the rolling sub-periods:

---Table 2---

3. Preliminary analysis

3.1 Test on the appropriateness of benchmark allocation

To begin with, we examine whether the 'true' benchmarks (the ones more appropriate for funds' Morningstar Global Category) provide a better fit than the self-declared prospectus benchmark. To estimate this, we use the R-squared from equations (3) and (4) as a proxy for the accuracy of the benchmark used:

$$R_{it} = \alpha_i + \beta_i R_{S\&P500} + e_{it} \quad (1)$$

$$R_{it} = \alpha_i + \beta_i R^{\text{'True' Benchmark}} + e_{it} \quad (2)$$

In this analysis we exclude the mutual funds that belong to the Large Cap blend category (460 funds, as per Table 1) for which the 'true' benchmark (S&P500) is the same as their prospectus benchmark. For the remaining 821 funds in the sample we estimate equation (1) and (2), over the 22 rolling windows. Figure 1 depicts the average R-squared across the funds in each sub-period obtained using the S&P as the benchmark (lower line) and 'true' benchmarks (upper line), as per equations (1) and (2).

---Figure 1---

The results confirm our notion that ‘true’ benchmarks are more suited for funds outside the Large Cap Blend category than the S&P500 index. The R-squared obtained using ‘true’ benchmarks for each sub-period and for the entire sample period is on average 10% higher, with peaks in 1999 and 2012, when the difference reached 14% and 11.5%. Given these results, the question that imposes itself is that of the impact of poorly suited benchmarks on the mutual fund performance and their ranking relative to other funds: do funds with a prospectus benchmark unsuitable for their investment style tend to outperform those benchmarks and whether they remain at the top of the fund rankings when the benchmark is swapped for the ‘true’ one.

3.2 Presence of alphas in passive benchmarks

The second issue we try to avoid in our assessment of performance ranks is that of the ‘closet-indexing’. For instance, if a ‘true’ benchmark (say, Russell 1000 Value) performs better than the self-designated benchmark chosen by a fund (S&P 500 here), the fund that belongs to that specific category (Large Cap Value in this example) is likely to outperform its self-reported benchmark (S&P 500), even if they are simply replicating their ‘true’ benchmark (Russell 1000 Value). That also means that such funds may rank higher relative to some other funds even though the fund managers exhibit no true skill.

To illustrate such bias inflicted by indices, in spirit of Costa and Jakob (2006), Chan, Dimmock, and Lakonishok (2009), Cremers, Petajisto and Zitzewitz (2012)⁴, we estimate standard Carhart four-factor alphas of both self-declared prospectus benchmark (S&P500) and the ‘true’ benchmarks in our sample:

$$R_{Benchmark,t} - R_f = \alpha_{Benchmark} + \beta_M(R_{M,t} - R_{F,t}) + \beta_{SMB}SMB_t + \beta_{HML}HML_t + \beta_{WML}WML_t + e_t \quad (3)$$

Where $R_{Benchmark,t}$ is the return on the (prospectus or ‘true’) benchmark index used; R_f is the US 1 month Treasury bill; $R_{M,t} - R_{F,t}$ is the market risk premium⁵; SMB and HML are size and value

⁴ who report non-zero alphas for passive benchmark indices

⁵ US market risk premium is defined as the value-weighted return of all CRSP firms incorporated in the US and listed on the NYSE, AMEX, or NASDAQ (R_m) minus one month US Treasury bill (R_f)

factors from Fama and French (1993) paper and WML is the Carhart (1997) momentum factor. $\alpha_{Benchmark}$ represents the four-factor (prospectus or ‘true’) benchmark alpha, i.e. the excess return of the benchmark unexplained by the four factors.

The four-factor Carhart alpha is calculated for the S&P500, Russel 1000 Growth, Russel 1000 Value, Russel Midcap and Russell 2000 over 36 monthly rolling periods, to obtain alphas from 1994 to 2016. The aforementioned benchmarks correspond to the five largest Morningstar Global categories in the data set and represent 80 percent of our fund sample (1,029 funds of a total 1,281). The remaining indices and their corresponding categories in our sample are not used for this analysis as the number of funds per category is not large enough resulting in some sub-periods featuring very few funds, jeopardising the objectivity of the results.

---Figure 2---

Figure 2 depicts the trend of annualized four-factor alphas (in bps) of the five indices. First, in line with previous studies (see for instance Mateus et al., 2017) the alphas of the five passive benchmarks are not zero. Specifically, the S&P500 and Russell 1000 Growth alphas more stable and tend to be more positive than those of the remaining indices analysed here. In the full sample period from January 1992 to February 2016, the S&P 500, Russell 1000 Growth and Russell Midcap indices all have positive four-factor alphas of 33.01, 74.93 and 60.17 basis points per year respectively; while the negative alphas of -12.58 and -197.01 basis points per year are obtained for the Russell 1000 Value and Russell 2000 index.

To obtain an indication of the magnitude of possible biases in fund performance evaluation by selecting an index not corresponding to funds’ risk profile and composition holdings, we calculate the difference between the Carhart alpha of the ‘true’ benchmark and the self-declared benchmark, S&P500, as per Figure 3. The difference is annualized and reported in basis points.

---Figure 3---

Figure 3 illustrates that S&P500 four-factor alphas differ from the remaining four indices corresponding to the global categories where most of our funds ‘reside’. For instance, in the sub-periods 1994-1996 and 1996-1998, the alpha for the self-declared prospectus benchmark S&P500 is positive but at least 100bps lower than the alpha for Russel 1000 Growth index.

This tendency of the ‘true’ benchmark alpha to be higher than the prospectus benchmark one is present in 20 out of 22 rolling windows in this study. Therefore, a mutual fund which is in Large Cap Growth category may take the benefit of the lower prospectus benchmark alpha relative to the ‘true’ one, more typical for its composition and risk. If such fund “beats” the prospectus benchmark, investors may view that as a vouch for managerial skill, whereas the fund may be simply replicating Russell 1000 Growth, not having any stock picking skill. Hence, its outperformance over prospectus benchmark should simply be attributed to a higher alpha of the ‘true’ benchmark, more appropriate for the given fund. However this is not the case for all the indices. Inverse situation can be noticed for Russel 2000, whose four-factor alpha is systematically lower than the S&P500 one.

To avoid the impact of these ‘closet indexers’, there is a need to look at the benchmark-adjusted performance of funds. This is particularly important to note when measuring funds ‘true’ performance and ‘true’ ranking relative to other funds both within the Morningstar Global Category peer group and overall. In the following section we present the methodology that appropriately adjusts performance and provides funds’ benchmark-adjusted alphas.

4. Performance and ranking methodology

To obtain unbiased alphas for funds, we apply Angelidis, Giamouridis and Tessaromatis (2013) adjustment, suggested in recent literature on performance measurement. The model is of interest to academics and investment professionals, as it adjusts the left hand side of the standard Carhart (1997) model by replacing the risk-adjusted return with the benchmark-adjusted return, i.e. the funds’ tracking error:

$$R_{i,t} - R_{Benchmark,t} = \alpha_i^* + \beta_{i1}^*(R_{M,t} - R_{F,t}) + \beta_{i2}^*SMB_t + \beta_{i3}^*HML_t + \beta_{i4}^*WML_t + e_i^* \quad (4)$$

where $R_{i,t} - R_{Benchmark,t}$ is the excess return of a mutual fund i over a benchmark in period t . As in equation (3) SMB and HML are size and value factors from Fama and French (1993) paper and WML is the Carhart (1997) momentum factor. α_i^* is the difference of the fund’s and benchmark’s Carhart alpha (AGT-adjusted alpha hereafter). Similarly, Beta ($\beta_{i1}^*, \beta_{i2}^*, \beta_{i3}^*, \beta_{i4}^*$) is

the difference between the fund's and benchmark's Carhart betas. All of the factor data is from Kenneth French's website⁶.

The AGT model, therefore, enables us to obtain AGT-adjusted four-factor alpha of a fund that accounts for the alpha of the benchmark. To assess the change in rankings when the benchmark changes from the prospectus benchmark to the 'true' category benchmark, the model will be used twice for each equity fund: with the S&P 500 as a benchmark and with the 'true' benchmark relevant for the Morningstar category a fund belongs to:

$$R_{i,t} - R_{S\&P500,t} = \alpha_i^{*S\&P500} + \beta_{i1}^{*S\&P500}(R_{M,t} - R_{F,t}) + \beta_{i2}^{*S\&P500}SMB_t + \beta_{i3}^{*S\&P500}HML_t + \beta_{i4}^{*S\&P500}WML_t + e_i^{*S\&P500} \quad (5)$$

$$R_{i,t} - R_{True,t} = \alpha_i^{*True} + \beta_{i1}^{*True}(R_{M,t} - R_{F,t}) + \beta_{i2}^{*True}SMB_t + \beta_{i3}^{*True}HML_t + \beta_{i4}^{*True}WML_t + e_i^* \quad (6)$$

Where $R_{S\&P500,t}$ and $R_{True,t}$ are the return of the S&P 500 and 'true' Morningstar Global category benchmark respectively, $\alpha_i^{*S\&P500}$ is the difference between the Carhart alpha of fund i and its prospectus benchmark, S&P 500, α_i^{*True} is the difference between the Carhart alpha of fund i and the 'true' benchmark; $\beta_{i1}^{*S\&P500}, \beta_{i2}^{*S\&P500}, \beta_{i3}^{*S\&P500}, \beta_{i4}^{*S\&P500}$ are fund i exposures to market risk, size, style and momentum factors beyond the exposure of S&P500 to those risks and $\beta_{i1}^{*True}, \beta_{i2}^{*True}, \beta_{i3}^{*True}, \beta_{i4}^{*True}$ are the fund i 's four-factor betas adjusted by those of the 'true' benchmark for fund i 's category. The rest is as per equation (4).

We estimate equations (5) and (6) for each fund and each of the 22 rolling sub-periods. In total, we estimate 9,393 AGT S&P adjusted alphas, $\alpha_i^{*S\&P500}$, and the same number of AGT 'True' benchmark adjusted alphas, α_i^{*True} . We rank the funds according to their $\alpha_i^{*S\&P500}$ in each of the 22 sub periods and split the funds in quartiles according to the performance. We do the same with α_i^{*True} to obtain the second set of fund quartile rankings.

⁶ Kenneth French's website: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

5. Performance Results

Table 3 shows the number of unique funds analysed in each period, the average AGT S&P adjusted alphas, $\alpha_i^{*S\&P500}$, and the average AGT ‘true’ benchmark adjusted alphas, α_i^{*True} , for each of the 22 overlapping periods from January 1992 to February 2016. All alphas are annualised averages across all categories, expressed in basis points. The table also reports the difference between the $\alpha_i^{*S\&P500}$ and the α_i^{*True} . In 15 over 22 periods the $\alpha_i^{*S\&P500}$ are higher than the alternative (column 6 Table 3), implying that using S&P500 as a target instead of a more appropriate benchmark enhances performance. In some periods such as 2000-2002, 2003-2005 and 2006-2008 $\alpha_i^{*S\&P500}$ is higher than the α_i^{*True} in at least 79% of funds. In periods such as 2005-2007 and 2007-2009, the average $\alpha_i^{*S\&P500}$ exceeds average α_i^{*True} by over 200 basis points. However, this trend is not as pronounced post financial crises: from 2009-2011 period onwards we find lower percentage of funds (e.g. 11% in 2011-2013) with an average $\alpha_i^{*S\&P500}$ higher than the alternative. In the full (non-overlapping) sample period, deploying the S&P 500 as the AGT adjustment instead of a ‘true’ category benchmark, overstates the performance for 61.68% of the funds. For the overall sample this difference in AGT adjusted alphas is significant at 1% level using Wilcoxon rank sum test ($Z = -5.326$).

---Table 3---

Although this evidence is pointing that using S&P500 as a benchmark in AGT model results in a better performance for a fund relative to the ‘true’ benchmark in most of the rolling sub-periods, we do not know whether this benefits more the funds at the top or at the bottom of performance ranks. One should not ignore the fact that there is still 38.8% of the funds in our sample that are worse off by indicating S&P500 as a prospectus benchmark. To further examine the issue of strategic benchmark choices, we investigate whether the fund rankings change considerably when the prospectus benchmark is replaced with a ‘true’ one.

6. The impact of benchmark choice on fund rankings

To conduct the analysis further we examine how the choice of benchmark may impact funds’ relative ranking: do winners tend to stay winners and do losers remain losers when the benchmark changes from the one disclosed in the prospectus (S&P 500) to the ‘true’

benchmark. Using the AGT adjusted alphas for each fund over 22 rolling periods, we split the funds into quartiles according to their performance in each period. Two sets of quartile rankings are constructed one based on AGT S&P500 adjusted alphas and one on AGT ‘true’ benchmark adjusted alphas. Quartile ranking is not done on a Morningstar global category basis as categories have small number of funds in a number of sub-periods. We construct quartiles using the funds in all the categories excluding those assigned to the Large Cap Blend Global category, as their ‘true’ benchmark is their prospectus benchmark, the S&P500 index.

We then examine the proportion/number of funds that change quartiles when the benchmark changes. Table 4 displays the number of funds in each quartile per rolling period, average annualised AGT alphas (in bps) adjusted for i) prospectus, S&P500 ($\alpha_i^{*S\&P500}$) and ii) ‘true’ category benchmark (α_i^{*True}) per each rolling period. The table also reports the difference between the two alphas, which signals the magnitude of a possible bias when inappropriate benchmark is used in performance assessment. The last column in Table 4 shows the percentage of funds that remain in the same performance quartile when S&P 500 index is replaced with the ‘true’ benchmark.

---Table 4---

In Panel A (Quartile 1) ‘true’ benchmarks provide a lower average AGT adjusted alpha in 12 out of 22 periods analysed, indicating that for 55% of the periods performance of winners estimated with S&P500 is overstated. Analogous tendency can be viewed for the Quartile 2, 3 and 4 (11, 13 and 11 out of 22 periods of lower average AGT ‘true’ benchmark adjusted alphas, α_i^{*True} , respectively).

More importantly, the average number of funds that remains in the top quarter over the years (70%), implies that 30% of the top performing funds drop in performance ranks and leave the quartile when the performance is adjusted with the ‘true’ benchmark. Overall, for the total sample period, there is on average 68bps advantage for Quartile 1 funds of using S&P 500 as the prospectus benchmark. Comparing this value to the equivalent average alpha difference in Panels B-D, it becomes evident that the top performing funds benefit most from the choice of prospectus benchmark. Panel D in fact suggests that Quartile 4 funds get penalised from inadequate benchmark selection. Thus, on average, close to 30% of funds move up in quartile

rankings when their performance is assessed against a ‘true’ global category benchmark. The average ‘true’ AGT adjusted alpha for the total period is 33 basis points higher than the one estimated with prospectus benchmark, leading us to conclude that these funds would be better off selecting a ‘true’ benchmark.

Quartile 2 and Quartile 3 funds (Panel B and C of Table 4) are of least interest to investors; the funds in these quartiles are neither the top funds investors look out for nor the ones at the bottom they are trying to avoid. However, we document that the results for both quartiles are similar: adjusting alphas with the ‘true’ benchmark changes, on average, the quartile ranking of 45% and 43% of funds from Quartile 2 and Quartile 3, respectively. Those movements can be in both directions – up to a higher or down to a lower ranked quartile, and in most of the cases there is an interchange between these two groups. The AGT alpha adjusted ‘true’ with the true benchmark is on average 28 (Quartile 2) and 26 (Quartile 3) basis points higher than the one adjusted with self-declared prospectus benchmark.

Therefore, inferences on mutual fund relative performance may be significantly biased when fund performance is evaluated in respect to unsuitable benchmark. To support our discussion, we plot the difference in average AGT adjusted alphas, $\alpha_i^{S\&P500} - \alpha_i^{True}$ in each ranking quartile and each period (column 5 from Table 4) in Figure 4.

---Figure 4 ---

The figure shows that the average AGT adjusted alphas for the Quartile 2 and 3 are almost identical irrespective of the benchmark. However, the performance of top funds is overestimated with the prospectus benchmark in over half of the rolling periods. The difference in AGT adjusted alphas for Quartile 1 funds reaches peaks of -241bps in 1993-1995, around -300 bps in 2006-2008 and 2008-2010, and a maximum of -460bps as in 2007-2009, in favour of alphas adjusted with the S&P rather than the ‘true’ benchmark. Even though top performing funds seem to take advantage of using S&P500 as their benchmark; there are also cases when performance of these funds benchmarking against prospectus benchmark could be undervalued, as in 2000-2002 and 2001-2003, but by a smaller margin. In contrast, our results show, that while benchmarking against prospectus benchmark is on the whole beneficial to winners, it negatively affects the performance of losers. The difference in AGT adjusted alphas

of Quartile 4 funds in some time periods, for instance 2011-2013 and 2012-2014, reaches 392 and 309 basis points, respectively, in favour of alphas adjusted with the ‘true’ benchmark.

Considering these findings, most of the funds that are potentially strategically selecting S&P 500 as the benchmark and benefiting from it are those in top performance quartile. They have on average 0.68% higher benchmark-adjusted alphas when that benchmark is the one given in the prospectus and nearly 30% of those funds lose the ‘winner’ status when the self-declared benchmark is substituted with a better suited one. In all other quartiles there is no clear advantage of using S&P 500 as a prospectus benchmark. Hence, the choice of the benchmark affects not only the inferences about a fund’s absolute performance, but it can also mislead investors about its relative performance. This leads us to conclude that any information in fund prospectus about the performance relative to the prospectus benchmark or relative to other funds should be treated with caution.

Conclusions

This paper analyses the impact of benchmark choice on US equity funds performance and gauges potential biases in absolute and relative performance evaluation stemming from the inaccurate prospectus benchmark selection. We re-visit the question of mismatch between the prospectus benchmark and fund objectives, raised in Sensoy (2009), and estimate the impact of such misclassification on fund performance and ranking. In contrast to the previous literature, our analysis accounts for non-zero benchmark alphas produced by standard pricing models, discussed in recent literature such as Chan, Dimmock, and Lakonishok (2009), Cremers, Petajisto and Zitzewitz (2012). Our sample includes net monthly returns of 1281 actively managed US equity mutual funds from January 1992 to February 2016, that report the S&P 500 index as their primary prospectus benchmark in the Morningstar database. We find that only 460 of those funds belong to the Large Cap blend Morningstar category, for which the S&P 500 would be the most suited benchmark. All other remaining funds fall across 21 other distinct Morningstar Global categories, some of which imply that fund risk profile and composition is very different from that of their prospectus benchmark. Naturally, we investigate whether the fund’s performance relative to the S&P 500 is better than when measured against what we consider their ‘true’ category benchmark. Regression of mutual fund returns on the returns of S&P500 and the ‘true’ Morningstar Global category benchmark a fund

belongs to, shows that 'true' category benchmarks are a better fit for our funds, having on average 10% higher R-squared.

Further, in our preliminary analysis, similar to Cremers et al. (2012) and Chinthalapati et al. (2017), we report non-zero alphas of passive benchmark indices in our sample. To eliminate the upward/downward biases in performance assessment caused by embedded benchmark alphas we apply Angelidis et al. (2013) method (AGT) that adjusts fund's alpha for benchmark's alpha, hence isolating manager's skill above that common to the benchmark. Performance for each fund is calculated against the prospectus and the 'true' benchmark, more appropriate for the Morningstar category a mutual fund belongs to. The sample period is split into 22 rolling overlapping windows, each being 36 month in length. Overall, for the total sample period we document higher AGT four-factors alphas estimated with S&P 500 as a benchmark versus those adjusted with the 'true' category benchmark. Fund performance estimated with the Carhart model for each mutual fund with 36 month period revealed that in 70 percent of the periods the average AGT alphas adjusted with S&P500 are higher and overestimate fund performance. Overall figures for the entire period show that 61.2% of the funds benefit from wrongly benchmarking their performance against S&P500 (prospectus benchmark). So that, the average AGT-adjusted alpha drops by 23 basis points when 'true' global categories benchmark is used.

Additional results show that inaccurate benchmark choice also influence relative performance assessment. Thus, 30% of top performing funds move their ranking position when their performance is adjusted with the 'true' benchmark. Surprisingly, the results show that the worst performing funds get penalised by stating inaccurate benchmark in their prospectus. In fact, close to 30% of losers move up their position when performance is estimated with the most suitable global category benchmark. The results indicate that the top quartile funds benefit most from the choice of prospectus benchmark. Thus, for instance, in 2007-2009 the difference in S&P 500 and 'true' benchmark adjusted alphas reached 460 bps in favour of using the prospectus benchmark. This leads us to conclude that strategic benchmark selection appears to be most likely in the funds at the top performance quartile, while we do not observe clear advantage of benchmark gaming in the remaining quartiles. Overall, we can highlight that the average alpha when the performance is adjusted with 'true' Global category benchmarks drops by 68bps in Quartile 1, falls 28bps and 25bps in Quartiles 2 and 3, and increases 33bps in Quartile 4 in the whole sample period.

These results show that appropriate benchmarking is essential for accurate performance evaluation, as inferences on both fund performance and performance ranking may change significantly when estimated against a ‘true’ benchmark instead of their self-declared prospectus benchmark. It is irrefutable in this paper that information disclosed by equity mutual funds regarding fund’s self-designated benchmark is by and large not accurate as prospectus benchmarks do not represent funds’ actual investment style in 2/3 of the funds in our sample. This study raises concerns that require attention of financial regulators and policy makers. New information disclosure requirements should be placed to provide more clarity for investors as to how the prospectus benchmark is selected. It also calls for investors to be more cautious when interpreting performance figures in fund prospectus. The paper can be extended to non-equity funds or funds where the benchmarking is ambiguous (such as hedge funds for instance).

References

- Angelidis, T., Giamouridis, D. and Tessaromatis, N., 2013. Revisiting mutual fund performance evaluation. *Journal of Banking & Finance*, 37(5), pp.1759-1776.
- Bams, D., Otten, R. and Ramezanifar, E., 2016. Investment Style Misclassification and Mutual Fund Performance, working paper.
- Carhart, M.M., 1997. On persistence in mutual fund performance. *The Journal of finance*, 52(1), pp.57-82.
- Chan, L.K., Dimmock, S.G. and Lakonishok, J., 2009. Benchmarking money manager performance: Issues and evidence. *Review of Financial Studies*, 22(11), pp.4553-4599.
- Chinthalapati, V.L., Mateus, C. and Todorovic, N., 2017. Alphas in disguise: A new approach to uncovering them, *International Journal of Finance and Economics*, forthcoming
- Costa, B. and K. Jakob, 2006. Do Stock indexes Have Abnormal Performance?, *The Journal of Performance Measurement* 11 (1), pp. 8-18
- Cremers, K.M. and Petajisto, A., 2009. How active is your fund manager? A new measure that predicts performance. *Review of Financial Studies*, 22(9), pp.3329-3365.
- Cremers, M., Petajisto, A. and Zitzewitz, E., 2012. Should Benchmark Indices Have Alpha? Revisiting Performance Evaluation, *Critical Finance Review*, 2, pp.1-48
- Dan diBartolomeo and Witkowski, E., 1997. Mutual fund misclassification: Evidence based on style analysis. *Financial Analysts Journal*, pp.32-43.
- Elton, E., Gruber, M., Blake, C., 2003. Incentive fees and mutual funds. *Journal of Finance* 58, pp.779–804.
- Fung, W. and Hsieh, D.A., 2002. Hedge-fund benchmarks: Information content and biases. *Financial Analysts Journal*, pp.22-34.
- Goetzmann, W., Ingersoll, J., Spiegel, M. and Welch, I., 2007. Portfolio performance manipulation and manipulation-proof performance measures. *Review of Financial Studies*, 20(5), pp.1503-1546.
- Huang, J., Sialm, C. and Zhang, H., 2011. Risk shifting and mutual fund performance. *Review of Financial Studies*, 24(8), pp.2575-2616.
- Jiang, H., Verbeek, M. and Wang, Y., 2014. Information content when mutual funds deviate from benchmarks. *Management Science*, 60(8), pp.2038-2053.
- Kacperczyk, M., Sialm, C. and Zheng, L., 2008. Unobserved actions of mutual funds. *The Review of Financial Studies*, 21(6), pp.2379-2416.

Kim, M., Shukla, R. and Tomas, M., 2000. Mutual fund objective misclassification. *Journal of Economics and Business*, 52(4), pp.309-323.

Mateus, I.B., Mateus, C. and Todorovic, N., 2016. UK equity mutual fund alphas make a comeback. *International Review of Financial Analysis*, 44, pp.98-110.

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

Figure 1: Average R-squared of S&P500 and 'True' benchmark fit

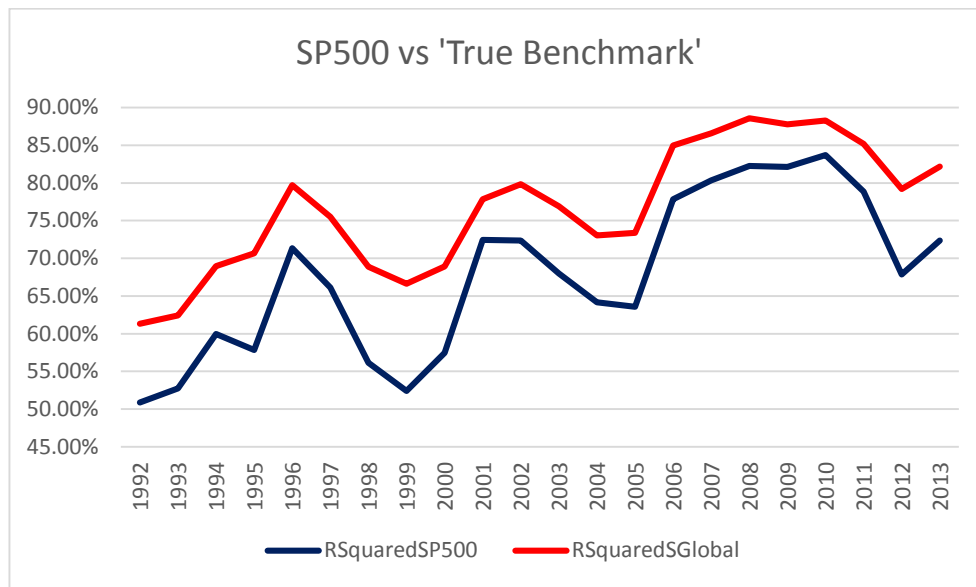


Figure 2: Four factor (Carhart) alphas of S&P 500 and selected ‘true’ benchmarks

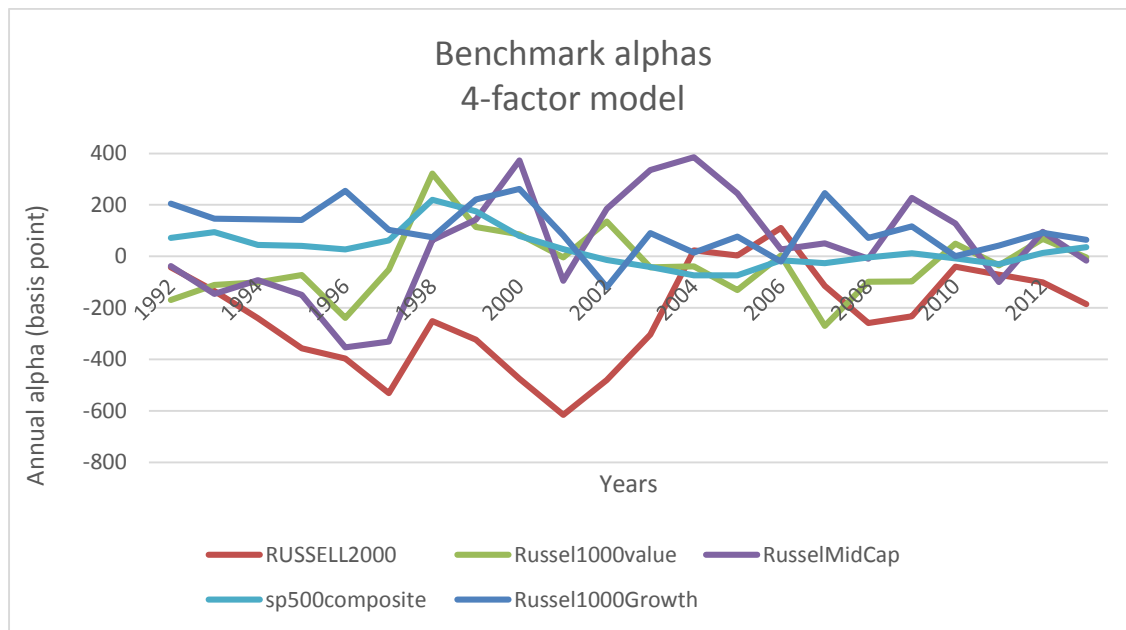


Figure 3 Differences between Carhart alphas of selected ‘true’ benchmarks and the S&P500

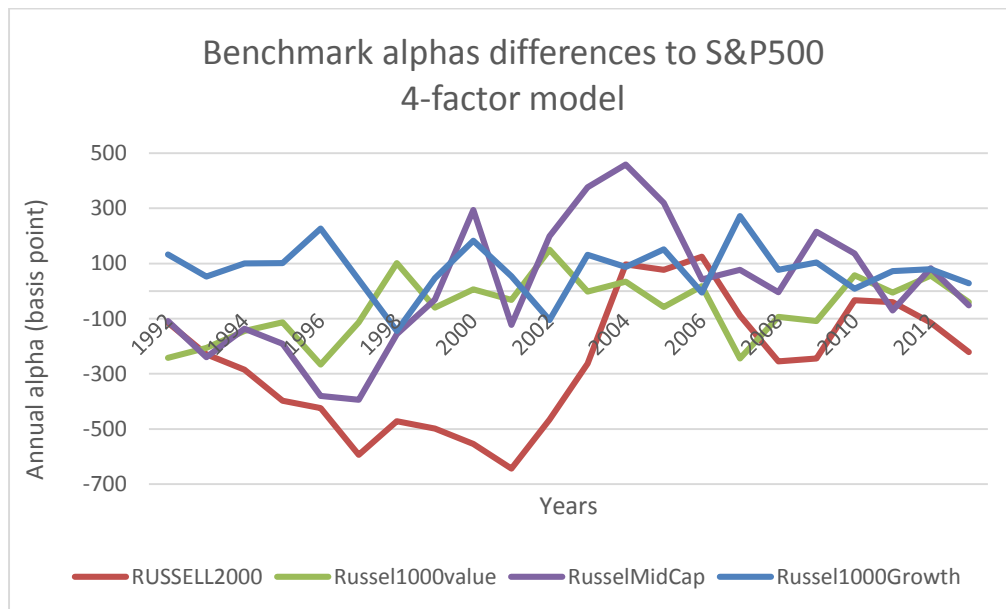


Figure 4: Difference between AGT S&P adjusted and AGT 'true' benchmark adjusted alphas

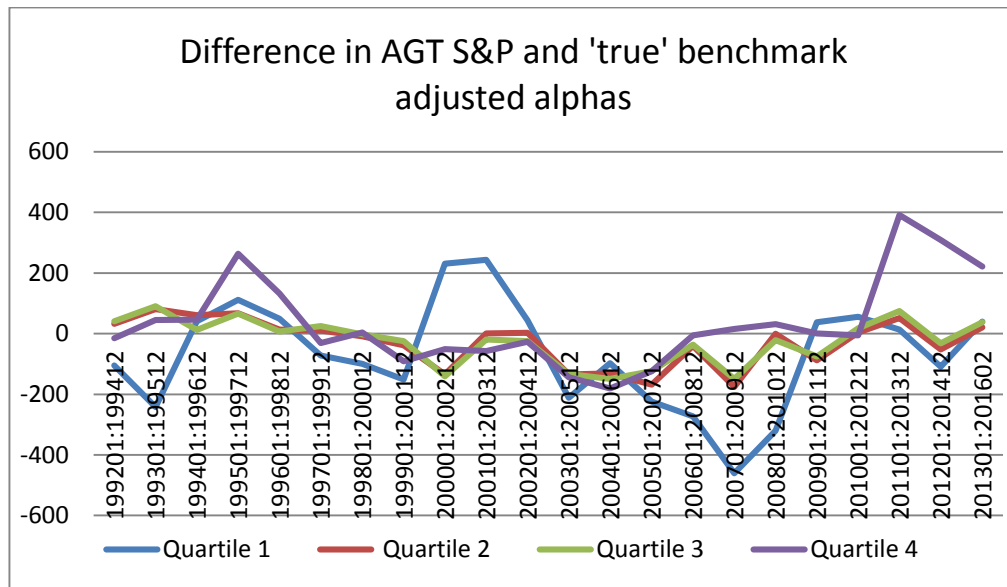


Table 1: Sample of ‘true’ benchmarks

The sample consists of 1,281 (212,122 monthly observations) long-only active US equity mutual funds from January 1992 to February 2016. For all funds the self-declared prospectus benchmark is the S&P500. Table below shows the Morningstar Global Category our funds belong to, the suitable benchmark for the category, the number of funds in the category and number of monthly observations per category (all benchmarks are total return and in USD).

Global Category	Suitable Benchmark	# Funds	# Monthly Observations
US Large Cap Blend	S&P 500	460	73,493
US Large Cap Growth	RUSSELL 1000 GROWTH	290	48,393
US Large Cap Value	RUSSELL 1000 VALUE	127	21,160
US Mid Cap	RUSSELL MIDCAP	112	17,332
Technology Sector Equity	S&P500 ES INFO TECHNOLOGY	54	9,092
US Small Cap	RUSSELL 2000	40	5,611
Healthcare Equity	S&P500 ES HEALTH CARE	32	5,554
Real Estate Equity	S&P500 DIVERSIFIED REIT'S	24	2,279
Global Equity	MSCI WORLD	22	3,392
Financial Sectors Equity	S&P500 DIVERSIFIED FINANCIALS	19	4,162
Energy Sector	S&P500 ENERGY IG	16	3,198
Precious Metals Sector Equity	S&P GSCI Precious Metal Tot. Ret.	16	4,196
Utilities Sector	S&P500 ES UTILITIES	14	3,293
Natural Resources Equity	S&P GSSI NORTH AMER. NAT.RES.SECTOR	13	2,400
Consumer Goods and Services	S&P500 ES CONSUMER DISCRETIONARY	12	3,159
Industrials Equity	S&P500 ES INDUSTRIALS	8	2,124
Communications Equity	S&P500 COMM. EQUIPMENT	8	1,268
Global Equity Large Cap	MSCI EAFE	7	1,174
Emerging Markets Equity	MSCI EM	2	324
Other Equity (Emerging Europe)	MSCI EM EUROPE	1	227
Europe Large Cap Equity	MSCI EUROPE	1	82
Asia Equity ex Japan	MSCI AC ASIA PAC EX JP	1	58
Japan Equity	MSCI JAPAN	1	61
Greater China	MSCI GOLDEN DRAGON	1	90
Total:		1,281	212,122

Table 2: Sample funds with more than 36 monthly observations

Table reports the number of funds and monthly observations for each of the 36 months rolling windows. The minimum data requirement is for funds to have at least 36 months of continuous observations. The #Funds represents the number of (non-unique) funds with available data in each period.

<i>Period</i>	<i># Funds</i>	<i># Monthly Observations</i>	<i>Period</i>	<i># Funds</i>	<i># Monthly Observations</i>
<i>199201:199412</i>	409	12,508	<i>200301:200512</i>	1,034	32,887
<i>199301:199512</i>	451	14,042	<i>200401:200612</i>	1,070	33,361
<i>199401:199612</i>	527	15,740	<i>200501:200712</i>	1,066	33,956
<i>199501:199712</i>	600	17,860	<i>200601:200812</i>	1,054	34,366
<i>199601:199812</i>	681	20,463	<i>200701:200912</i>	1,057	33,663
<i>199701:199912</i>	771	23,364	<i>200801:201012</i>	1,039	32,453
<i>199801:200012</i>	865	26,305	<i>200901:201112</i>	975	30,906
<i>199901:200112</i>	919	28,916	<i>201001:201212</i>	895	29,500
<i>200001:200212</i>	955	30,874	<i>201101:201312</i>	855	27,929
<i>200101:200312</i>	980	32,085	<i>201201:201412</i>	789	26,519
<i>200201:200412</i>	997	32,640	<i>201301:201602</i>	751	26,573
<i>Overall: 199201:201602</i>				1,281	211,855

Table 3: Comparison of average AGT S&P adjusted alphas and average AGT ‘true’ benchmark adjusted alphas

The table reports comparison of alphas from the following two regressions:

$R_{i,t} - R_{S\&P500,t} = \alpha_i^{*S\&P500} + \beta_{i1}^{*S\&P500}(R_{M,t} - R_{F,t}) + \beta_{i2}^{*S\&P500}SMB_t + \beta_{i3}^{*S\&P500}HML_t + \beta_{i4}^{*S\&P500}WML_t + e_i^{*S\&P500}$ and $R_{i,t} - R_{True,t} = \alpha_i^{*True} + \beta_{i1}^{*True}(R_{M,t} - R_{F,t}) + \beta_{i2}^{*True}SMB_t + \beta_{i3}^{*True}HML_t + \beta_{i4}^{*True}WML_t + e_i^{*True}$. $R_{S\&P500,t}$ and $R_{True,t}$ are the return of the S&P 500 and ‘true’ Morningstar Global category benchmark respectively, $\alpha_i^{*S\&P500}$ is the difference between the Carhart alpha of fund i and its prospectus benchmark, S&P 500, α_i^{*True} is the difference between the Carhart alpha of fund i and the ‘true’ benchmark; $\beta_{i1}^{*S\&P500}, \beta_{i2}^{*S\&P500}, \beta_{i3}^{*S\&P500}, \beta_{i4}^{*S\&P500}$ are fund i exposures to market risk, size, style and momentum factors beyond the exposure of S&P500 to those risks and $\beta_{i1}^{*True}, \beta_{i2}^{*True}, \beta_{i3}^{*True}, \beta_{i4}^{*True}$ are the fund i ’s four-factor betas adjusted by those of the ‘true’ benchmark for fund i ’s category. $R_{M,t} - R_{F,t}$ is the market risk premium; SMB and HML are size and value factors from Fama and French (1993) paper and WML is the Carhart (1997) momentum factor. Alphas and the difference in alphas are annualized and given in bps.

<i>Period</i>	<i># of funds</i>	<i>Average $\alpha_i^{*S\&P500}$ (bp)</i>	<i>Average α_i^{*True} (bp)</i>	<i>Average difference $\alpha_i^{*True} - \alpha_i^{*S\&P500}$ (bp)</i>	<i>Better $\alpha_i^{*S\&P500}$ #/%</i>	<i>Better α_i^{*True} #/%</i>
199201:199412	192	43	4	-39	110/58%	81/42%
199301:199512	218	28	-12	-40	121/56%	97/44%
199401:199612	245	-208	-151	57	113/46%	132/54%
199501:199712	275	-356	-217	139	124/45%	151/55%
199601:199812	299	-274	-261	14	143/48%	156/52%
199701:199912	344	-318	-363	-45	165/48%	179/52%
199801:200012	384	341	214	-127	261/68%	123/32%
199901:200112	433	279	198	-81	310/72%	123/28%
200001:200212	476	47	-14	-61	392/82%	84/18%
200101:200312	526	-244	-217	27	311/59%	215/41%
200201:200412	534	-154	-156	-2	287/54%	247/46%
200301:200512	514	29	-111	-140	426/83%	88/17%
200401:200612	524	8	-85	-93	402/77%	122/23%
200501:200712	514	333	78	-255	384/75%	130/25%
200601:200812	513	113	-17	-130	405/79%	108/21%
200701:200912	513	228	2	-226	382/75%	131/25%
200801:201012	506	87	-32	-119	345/68%	161/32%
200901:201112	490	-9	120	129	330/67%	160/33%
201001:201212	487	-225	-178	47	201/41%	286/59%
201101:201312	484	-361	-154	207	55/11%	429/89%
201201:201412	473	-313	-183	130	200/42%	273/58%
201301:201602	449	-330	-231	99	280/62%	169/38%
Overall					5,747/61.2%	3,645/38.80%
Overall*					2,071/61.18%	1,278/38.16%

*no overlapping period

Table 4: Difference in alphas per quartile and change of quartile ranks

Panels A-D report results for Quartile 1(top) - 4 (bottom) respectively. All panels show the number of funds and comparison of AGT adjusted alphas, when S&P 500 is used as a benchmark ($\alpha_i^{S\&P500}$ from eq(5)) and when 'true' benchmark is used (α_i^{True} from equation (6)). Alphas and the difference in alphas are annualised and given in basis points. The last column shows percentage of funds that remains in the same quartile when the benchmark is changed from the S&P500 to the 'true' benchmark. In the last row, the 'average' represents the average across the periods and across the funds.

Panel A: Quartile 1 (Carhart model)					
<i>Period</i>	<i># of Funds</i>	<i>Average</i> $\alpha_i^{S\&P500}$ <i>(bp)</i>	<i>Average</i> α_i^{True} <i>(bp)</i>	<i>Average difference</i> $\alpha_i^{True} - \alpha_i^{S\&P500}$ <i>(bp)</i>	<i>% Funds</i> <i>remaining in</i> <i>Quartile 1</i>
199201:199412	48	724.622	619.908	-104.713	77.08
199301:199512	55	900.993	659.124	-241.869	78.18
199401:199612	61	509.698	550.235	40.537	72.13
199501:199712	69	406.961	518.888	111.927	71.01
199601:199812	75	401.890	450.835	48.945	50.67
199701:199912	86	710.099	637.692	-72.407	75.58
199801:200012	96	1513.924	1414.484	-99.440	68.75
199901:200112	108	1431.359	1279.641	-151.718	76.85
200001:200212	119	1079.855	1310.632	230.777	64.71
200101:200312	132	744.157	987.541	243.384	68.94
200201:200412	134	477.240	521.377	44.137	65.67
200301:200512	129	649.942	438.095	-211.847	72.09
200401:200612	131	643.598	545.615	-97.982	78.63
200501:200712	129	667.274	442.562	-224.712	68.22
200601:200812	128	675.317	402.982	-272.335	75.00
200701:200912	128	1049.981	589.334	-460.646	58.59
200801:201012	127	876.886	556.633	-320.253	77.95
200901:201112	123	586.092	623.681	37.589	61.79
201001:201212	122	390.537	446.654	56.1171	68.03
201101:201312	121	406.372	420.2614	13.889	71.07
201201:201412	118	496.183	386.647	-109.535	69.49
201301:201602	112	234.353	274.290	39.937	71.43
Average				-68.19	70.09

Panel B: Quartile 2 (Carhart model)					
<i>Period</i>	<i># of Funds</i>	<i>Average</i> $\alpha_i^{S\&P500}$ (bp)	<i>Average</i> α_i^{True} (bp)	<i>Average difference</i> $\alpha_i^{True} - \alpha_i^{S\&P500}$ (bp)	<i>% Funds</i> <i>remaining in</i> <i>Quartile 2</i>
199201:199412	48	66.766	98.823	32.057	43.75
199301:199512	54	20.647	101.976	81.329	53.70
199401:199612	61	-64.974	-3.744	61.229	52.46
199501:199712	69	-47.359	19.710	67.069	44.93
199601:199812	75	-83.858	-70.943	12.916	9.33
199701:199912	86	-103.845	-93.813	10.033	61.63
199801:200012	96	285.842	277.472	-8.370	54.17
199901:200112	108	284.427	246.554	-37.873	64.81
200001:200212	119	261.234	126.598	-134.636	64.71
200101:200312	131	-93.440	-93.051	0.389	64.89
200201:200412	133	-22.030	-19.082	2.948	45.11
200301:200512	128	117.070	-17.686	-134.756	49.22
200401:200612	131	197.198	66.072	-131.125	54.96
200501:200712	128	242.302	73.913	-168.389	57.03
200601:200812	128	84.538	44.437	-40.100	71.09
200701:200912	128	261.924	86.677	-175.248	32.03
200801:201012	126	47.480	46.723	-0.757	65.87
200901:201112	122	22.702	-65.162	-87.864	63.11
201001:201212	122	-17.213	-13.158	4.055	61.48
201101:201312	121	-37.196	13.985	51.181	71.07
201201:201412	118	-20.476	-72.422	-51.946	67.80
201301:201602	112	-121.013	-99.772	21.241	67.86
Average				-28.48	55.50

Panel C: Quartile 3 (Carhart model)					
<i>Period</i>	<i># of Funds</i>	<i>Average</i> $\alpha_i^{*S\&P500}$ (bp)	<i>Average</i> α_i^{*True} (bp)	<i>Average difference</i> $\alpha_i^{*True} - \alpha_i^{*S\&P500}$ (bp)	<i>% Funds</i> <i>remaining in</i> <i>Quartile 3</i>
199201:199412	47	-187.715	-147.318	40.397	38.30
199301:199512	54	-224.299	-133.920	90.379	51.85
199401:199612	62	-306.194	-294.377	11.817	50.00
199501:199712	68	-352.285	-286.392	65.893	52.94
199601:199812	74	-357.022	-350.239	6.783	25.68
199701:199912	86	-402.512	-377.958	24.555	66.28
199801:200012	96	-98.842	-100.839	-1.997	55.21
199901:200112	109	-85.088	-109.625	-24.537	75.23
200001:200212	119	-119.334	-260.565	-141.231	70.59
200101:200312	131	-446.355	-465.370	-19.015	64.12
200201:200412	133	-261.305	-285.984	-24.679	51.88
200301:200512	128	-105.933	-239.287	-133.353	49.22
200401:200612	131	-20.324	-169.740	-149.416	54.96
200501:200712	128	7.973	-116.063	-124.036	51.56
200601:200812	129	-129.902	-165.936	-36.034	78.29
200701:200912	129	19.864	-129.194	-149.058	34.11
200801:201012	126	-157.862	-179.444	-21.581	66.67
200901:201112	122	-259.939	-334.760	-74.821	55.74
201001:201212	121	-251.755	-234.360	17.395	67.77
201101:201312	121	-279.317	-205.105	74.211	66.94
201201:201412	119	-241.206	-273.286	-32.080	69.75
201301:201602	113	-320.076	-282.608	37.468	64.60
Average				-25.59	57.35

Panel D: Quartile 4(Carhart model)					
<i>Period</i>	<i># of Funds</i>	<i>Average</i> $\alpha_i^{*S\&P500}$ (bp)	<i>Average</i> α_i^{*True} (bp)	<i>Average difference</i> $\alpha_i^{*True} - \alpha_i^{*S\&P500}$ (bp)	<i>% Funds</i> <i>remaining in</i> <i>Quartile 4</i>
199201:199412	48	-680.588	-695.340	-14.752	72.92
199301:199512	55	-780.702	-735.892	44.809	69.09
199401:199612	61	-930.347	-884.578	45.769	67.21
199501:199712	69	-1436.729	-1173.304	263.425	69.57
199601:199812	75	-1073.939	-941.362	132.577	64.00
199701:199912	86	-1218.210	-1248.937	-30.727	80.23
199801:200012	96	-743.990	-740.403	3.588	68.75
199901:200112	108	-709.590	-800.757	-91.167	76.85
200001:200212	119	-839.593	-890.372	-50.779	68.91
200101:200312	132	-1038.163	-1096.031	-57.868	77.27
200201:200412	134	-770.944	-797.378	-26.434	79.10
200301:200512	129	-492.279	-635.970	-143.691	68.22
200401:200612	131	-445.152	-626.243	-181.091	78.63
200501:200712	129	-352.865	-477.177	-124.312	68.99
200601:200812	128	-517.257	-523.115	-5.858	80.47
200701:200912	128	-495.099	-480.083	15.015	53.91
200801:201012	127	-651.181	-620.318	30.863	81.10
200901:201112	123	-778.481	-777.586	0.895	65.04
201001:201212	122	-830.054	-835.978	-5.924	80.33
201101:201312	121	-1193.792	-802.262	391.529	69.42
201201:201412	118	-1137.554	-828.766	308.788	72.88
201301:201602	112	-832.393	-611.258	221.135	74.11
Average				32.99	72.14