Searching for mutual fund winners? The strategy is to outbid both, the benchmark and the peer group

Cesario Mateus

Aalborg University Business School, Aalborg University
Aalborg, Denmark
cmateus@business.aau.dk

Irina Mateus
Aalborg University Business School, Aalborg University
Aalborg, Denmark
imateus@business.aau.dk

Natasa Todorovic
Bayes Business School, City, University of London
London, UK
natasha.todorovic.1@city.ac.uk

Abstract

Standard Fama-French-Carhart models define 'winners' as funds that generate the highest excess returns given the factor risks involved; however, they do not provide information on whether such winners are outperforming their prospectus benchmark or their peer group. In addition, existing literature relying on these models, by and large, does not find evidence of persistence in performance. In this paper, we propose a two-stage procedure that allows investors to select "true" winners(losers) which generate the highest factor-risk-adjusted performance relative to the benchmark *and* the peer group simultaneously. Utilizing both adjustments at the same time results in a strong predictive ability, leading to a selection of funds that persist in performance. Our true winner funds have statistically significant superior benchmark-adjusted alphas, peer group adjusted alphas and Sharpe ratios one year ahead, which are significantly different from those generated by the true loser funds. The results are robust to extended investment horizon, and alpha estimation method, and they are not driven by outliers, size of fund-sorts, or any particular period within our sample.

Keywords: US Equity Mutual Funds, Benchmark-adjusted alphas, Peer-group-adjusted alphas, Performance ranking

JEL: G11, G12, G2

1. Introduction

The definition of winner funds' in the mutual funds literature is typically based on the highest risk-adjusted performance, such as the highest Sharpe ratios or Fama-French/Carhart three/four-factor alphas; with some studies considering as winners those funds with the highest market-adjusted returns (a non-risk adjusted performance measure, see for instance Fletcher and Forbes, 2002). In practice, though, a performance of a mutual fund is judged against its self-reported (prospectus) benchmark and against other similar funds comparable in terms of investment objectives. However, the fact that a fund is classified as outperforming relative to its self-reported benchmark does not tell us much about its performance relative to its peer group. Similarly, a fund that is at the top of the peer group may not necessarily outperform its benchmark. Undoubtedly, it is of importance to investors to identify true, unbiased winner funds. In this paper, we contribute to the mutual fund performance evaluation literature by demonstrating that benchmark-adjusted and peer-group-adjusted performance measurements should be used *jointly*, not exclusively. A true winner in our study is defined as a fund that, on a factor-risk-adjusted basis, outperforms the benchmark *and* the peer group at the same time. We show how those funds can be identified and prove that true winners (losers) that outperform (underperform) the benchmark *and* the peer group simultaneously exhibit persistence in performance.

During the last three decades, significant research has been focused on determining whether mutual funds can create value and outperform their corresponding benchmarks. The evidence on US mutual fund performance based on the standard Fama-French-Carhart models is mixed, where some studies suggest significant underperformance (see for instance Carhart, 1997, Fama and French 2010), while others find some evidence of outperformance (e.g. Wermers, 2000, Kosowsky et al., 2006). Looking at the fund performance by investment styles, Davies (2001), finds no evidence of outperformance of any particular style, and in fact documents an underperformance of 2.7% p.a. of value funds, based on the standard three-factor model. The international evidence of mutual fund performance is also mixed (see for instance Ferreira et al., 2013, Filip and Rogala, 2015, Gallagher, 2003, and Ding et al., 2015 for instance). Further, the evidence on mutual fund persistence in performance based on these standard models is also mixed, where persistence is found among losers but not pervasively among winner funds (see for instance Carhart, 1997, Elton et al., 1996, Brown and Goetzmann, 1995, among others). For a detailed review of mutual fund performance and persistence in performance literature see Cuthbertson et al., 2010, Mateus et al., 2019c and Cremers et al., 2019).

Our contribution is within the new line of research in mutual fund performance and persistence in performance, based on the work of Cremers et al., (2012). They provide evidence that even passive benchmarks generate non-zero alphas when regressed on the Carhart four-factor model, and point to the questionable arbitrary nature of Fama-French factor construction which leads to disproportionate weights assigned to value and small-cap stocks (see also Matallin-Saez, 2007; Elton and Gruber 2020). The new strand of academic literature (Mateus et al., 2016, 2019a, 2019b) claims that augmented versions of the Carhart model can be utilised to estimate the most accurate mutual fund's performance accounted for the embedded passive benchmark alphas and identifies three methodologies that help on this matter: 1) Angelidis, Giamouridis and Tessaromatis (2013) eliminates non-zero benchmark alphas by amending the left-hand side of the Carhart four-factor model, thus substituting the risk-adjusted return with the benchmark-adjusted return; 2) Chinthalapati et al., (2017) provides a mathematical solution to the problem of disproportionate weights and calculates minor fixed adjustments that should be added to the time series of the Carhart's factors to ensure a zero alpha for any self-designated benchmark index of a mutual fund (leaving the factor loadings, R-squared and other model parameters unchanged)[1]; 3) Hunter et al. (2014) approach suggests amending the Carhart model with an APB (Active Peer Benchmark) adjustment factor that mitigates the issue of passive benchmark mispricing within the peer group (Elton and Gruber, 2020) and the issue of mismatched reference benchmarks (see for instance Sensoy 2009, Bams, et al., 2015, Cremers and Petajisto, 2009, Elton, Gruber, and Blake, 2003).

Our work differs from the studies in this area whose focus was on examining the presence of non-zero alphas in passive indices (Cremers et al., 2012, Mateus et al. 2016), correcting for benchmarks' non-zero alphas by assessing funds' benchmark-adjusted performance *only* (Angelidis et al., 2013, Chinthalapati et al., 2017, Mateus et al., 2016) or funds' peer-group-adjusted performance *only* (Hunter et al. 2014, Mateus et al. 2019b). In this paper, we provide a method that unifies the approaches from this strand of literature. Note that Angelidis, Giamouridis and Tessaromatis (2013) method (AGT hereafter) gives fund's benchmark-adjusted performance relative to a self-reported benchmark but does not reveal how a fund

[1] As per Mateus et al. (2019c), both methodologies in 1) and 2) provide qualitatively and quantitatively similar results, with method 1) by Angelidis, Giamouridis and Thesarromatis (2013) approach being more commonly tested in the literature, hence adopted in this paper.

fares against its peer-group. On the other hand, Hunter et al. (2014) method deploy Active Peer Benchmark that helps select outperformers by accounting for commonalities within a group (we refer to the method as APB hereafter). However, it is possible that a fund at the top of the peer-group still underperforms its benchmark, just less so than the other funds in the group. We acknowledge from the previous literature that due to the presence of non-zero alphas in benchmark indices, one should adopt a method for determining unbiased mutual fund performance; and that funds' benchmark choice should be taken with caution (Sensoy, 2009, Mateus et al. 2019a). Taking these points into account, we contribute to the mutual fund persistence in performance literature by demonstrating in this study that in order to identify the 'true' unbiased winners (losers) across active mutual funds, the benchmark-adjusted and peer-group adjusted methods for assessing mutual fund performance should not be used independently – but jointly.

Our two-stage (benchmark adjusted and peer group-adjusted) performance assessment is based on the sample of 989 active US equity mutual funds collected from the Morningstar database for the period from 1992 to 2015. For the peer performance assessment, we construct an equally weighted portfolio of funds within a peer group (defined by a fund's Morningstar category). To eliminate a bias inflicted by possible mismatched self-reported benchmarks, for each Morningstar category has been selected a US index that best matches the category description and is widely used for benchmarking purposes. For instance, the S&P500 index is selected for the Large Cap Blend category, Russell 1000 Growth for the Large Cap Growth category etc.

Our results demonstrate that when AGT and APB methodologies are applied together (the combined AGT&APB procedure hereafter) 40 percent of winners (losers) selected based on only AGT or APB are eliminated. This indicates that those excluded funds do not comply with the definition of 'true' winners (losers) as they do not satisfy both criteria simultaneously: 1) reference benchmark outperformance (underperformance) and 2) the location in the top (bottom) quartiles among peers. Our empirical evidence shows that the winners (losers) identified with *both* AGT and APB methodologies continue to exhibit superior (inferior) performance in the years that follow. Most importantly, we provide evidence that the top (bottom) quartile funds selected based on *both* AGT and APB alphas, provide investors with better (worse) Sharpe ratios, AGT alphas and APB alphas up to three years ahead, than funds in the top (bottom) quartiles formed based on AGT or APB alpha solely. All the results stand when the performance is gauged with excess returns. Various robustness tests confirm our results.

Our paper provides a significant academic contribution to the current evidence on mutual fund performance relative to the benchmark and the peer group and confirms the existence of mutual funds' performance persistence. It is of high importance to financial practitioners and investors looking for an unbiased approach that enables them to select the 'true' winners with a consecutive persistent performance up to three years ahead. Our requirements to identify the 'true' winners are important for successful investment decision-making and may potentially be used in further research examining international mutual fund performance.

The paper is organised as follows: Section 2 describes the data and methodology, Section 3 provides an analysis of results and Section 4 concludes the paper.

2. Data and Methodology

2.1. Data

Our initial sample comprises 1,281 US active equity funds that report the S&P500 index as their primary prospectus benchmark. The funds in the sample span across all Morningstar global categories[2], however, we impose a requirement of a minimum of 10 funds per performance quartile in each category, due to our methodological procedure described in the next section. Therefore, our final sample contains mutual funds that belong to four categories, namely Large Cap Blend, Large Cap Growth, Large Cap Value, and Mid Cap[3] with a total sample of 989 US active equity funds and 160,378 monthly observations. Given that all funds report S&P500 index (a proxy for the large-cap blend style) as the primary benchmark, one would expect that all funds in our sample would fall into the Large Cap Blend category, which is not the case. To avoid the issue of mismatched or strategically chosen prospectus benchmark, discussed in recent literature (Sensoy 2009, Cremers et al., 2022), we followed the procedure of Mateus et. al., (2019a) and allocated to funds in each Morningstar category the US benchmark index most closely matching the description of the category. Hence, we paired the US Large Cap Blend category with the S&P500, Large Cap Growth with Russell 1000 Growth, Large Cap Value with Russell 1000 Value, and Mid Cap category with Russell Mid Cap Index.

The net monthly returns data of the funds (inclusive of dividends) and total returns of benchmark indices are collected from Morningstar. Our sample spans from January 1992 to December 2015 and is free from survivorship bias.

---Insert Table 1 here---

2.2. The application of the combined models

Academic evidence shows that non-zero passive benchmark alphas yielded by the standard Carhart model indicate biased under(out)-performance of active equity mutual funds (Cremers 2012, Chinthalapati et al., 2017, Elton and Gruber 2020). Such under(out)-performance is more pronounced during periods of adverse macroeconomic and market conditions (Chinthalapati et al., 2017). The model of Angelidis et al. (2013) allows investors to eliminate the bias driven by the Carhart portfolio construction. However, in addition to Mateus et al., (2019a) we claim

[2] Morningstar categorises US funds according to their portfolio holdings into five 'global categories' (Large Cap Value, Large Cap Growth, Large Cap Blend, Mid Cap and Small cap).

[3] The Small Cap fund category is not included as they do not satisfy our minimum funds requirement per performance quartile, as outlined in the methodology.

that to identify the definite winners a relative adjustment by the peer group is required, since funds may outbid the benchmark but not necessarily be the winners across their peer group and vice versa. Therefore we hypothesise that the application of both models the AGT and the APB will eliminate a part of the winners suggested by the stand-alone AGT or APB models. In this paper, we argue that the combined application of both AGT and APB models will lead to a more accurate selection of the "true" winner and loser funds with persistent performance[4].

Our proposed "two-step" approach is as follows. For each fund, we estimate the AGT alpha, following Angelidis et al. (2013) model. Thus, we replace the risk-free rate on the left-hand side of the standard Carhart (1997) equation with the return of the benchmarks as below:

$$R_{i,t} - R_{Benchmark,t} = \alpha_i^* + \beta_{i1}^*(R_{M,t} - R_{F,t}) + \beta_{i2}^* SMB_t + \beta_{i3}^* HML_t + \beta_{i4}^* WML_t + e_i^* \qquad (1)$$

where $R_{i,t} - R_{Benchmark,t}$ is the benchmark-adjusted return of a mutual fund i in period t[5]. The $(R_{M,t} - R_{F,t})$ is the market risk premium defined as the value-weighted return of all CRSP firms incorporated in the US and listed on the NYSE, AMEX, or NASDAQ minus one month US Treasury bill. The SMB and HML are Fama and French (1993) size and value factors, while WML is the Carhart (1997) momentum factor[6]. α_i^* is the fund's benchmark-adjusted alpha; a positive (and significant) value implying that the fund has outperformed the benchmark and vice versa. $\beta_{i1}^*, \beta_{i2,}^* \beta_{i3,}^* \beta_{i4}^*$ represent the difference between the fund's and benchmark's Carhart betas, i.e. benchmark-adjusted betas.

According to Frazzini et al. (2016) and Mateus et al. (2019a), the choice of benchmark in the models that utilise them, such as AGT model in equation (1) in our case, matters. One of the characteristics of a good benchmark suggested by the CFA[7]guidelines is that it should be consistent with manager's investment style. In this paper, by selecting the benchmark indices closely matching investment style of funds given by Morningstar global category classification, rather than fund's self-reported primary benchmark, we believe that we minimise the benchmark mis-specification problem.

---

[4] In this paper we do not aim to search for any variables omitted from the standard factor models
[5] The benchmarks for each Morningstar global category are presented in the Table 1.
[6] All factor data is obtained from Kenneth French's website: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.
[7] https://www.cfauk.org/-/media/files/pdf/pdf/5-professionalism/3-research-and-position-papers/benchmarks-and-indices.pdf (Accessed: 22/10/2022)

To proceed with the APB methodology we consider each Morningstar global category as the peer group for funds in that category. Following Hunter et al. (2014), we construct the Active Peer-group Benchmark as the equally weighted portfolio of all the funds within the category. We regress the excess return of the Active Peer-group Benchmark ($R_{APB,i,t}$) against the standard Carhart four-factors to obtain the alpha (α_{APB}) and the residual ($e_{APB,t}$), which are then used as the adjustment in the Carhart model that enables us to account for commonalities of alphas within a peer group. Equations (2) and (3) illustrate this:

$$R_{APB,t} = \alpha_{APB} + \beta_{APB,M}(R_{M,t} - R_{F,t}) + \beta_{APB,SMB}SMB_t + \beta_{APB,HML}HML_t + \beta_{APB,WML}WML_t + e_{APB,t} \tag{2}$$

$$R_{i,t} = \alpha_{i,ADJ} + \beta_{i,M}(R_{M,t} - R_{F,t}) + \beta_{i,SMB}SMB_t + \beta_{i,HML}HML_t + \beta_{i,WML}WML_t + \beta_{i,ADJ}(\alpha_{APB} + e_{APB,t}) + \omega_{i,ADJ,t} \tag{3}$$

In equation (2), α_{APB} is the alpha of the Active Peer-group Benchmark, $e_{APB,t}$ is the APB residual. In equation (3), $\alpha_{APB} + e_{APB,t}$ is the adjustment factor and the new $\alpha_{i,ADJ}$ is the APB-adjusted alpha. The APB alpha accounts for performance that is a result of a fund manager pursuing a common strategy and average risk of the peer group and, thus, isolates the fund manager's skill beyond that of the peer group. $(R_{M,t} - R_{F,t})$, SMB, HML and WML factors in equations (2) and (3) are defined as per equation (1).

2.3. Top and bottom quartile funds and their performance

The AGT and APB-adjusted alphas, described in Section 2.2., are estimated over 21 rolling windows of 36-months each[8], starting from January 1992 - December 1994 and ending with January 2012 - December 2014 window. In each of the rolling windows, we group the funds within each peer group into four performance quartiles according to estimated AGT and APB alpha separately. For this analysis, we set a requirement of a minimum of 10 funds within the performance quartile in each rolling period.

We argue here that not all the funds that are in the top quartile based on the AGT alpha will outperform their peer group as well as not all the funds in the top quartile relative to their peer group (top APB alphas) will outperform their benchmark. Hence, we claim that the joint

[8] We require minimum 30 months of returns in each of the rolling windows.

application of AGT and APB alphas will identify more defined winners and will eliminate part of the winners suggested by the AGT or APB models only. To explore this, we form performance quartiles according to *both* AGT and APB alphas with the expectation that the funds in the top (bottom) quartile formed in this manner will generate the best (worst) performance in the following year.

We gauge the performance of funds in the top (bottom) quartile one-year post-ranking by Sharpe ratios, defined as the excess return of a fund over the one-month US T-bill, per unit of the standard deviation of a fund. We require a fund to have a minimum of 6 monthly returns data following their assignment to quartiles to be included in this predictive analysis. For robustness, in addition to assessing performance through average Sharpe ratios in the top and bottom quartile of funds one-year post-ranking, we also assess their AGT and APB alphas one-year post-ranking (where appropriate) and their excess returns.

Table 2 shows the total number of funds in the top and bottom quartile with the available minimum of 6 months of data in the year following the ranking according to each model: AGT, APB and when the two models are used jointly.

---Insert Table 2 here---

Our results confirm our prediction and show that the combined application of AGT and APB models reduces the number of the top and bottom quartile finds, thus providing a more trimmed selection of the winners and losers. For instance, in the Large Cap Growth category where a suitable index would be Russell 1000 Growth, there are 695 funds in the top quartile when funds are ranked by their benchmark-adjusted alphas (AGT) and 717 funds are in the top quartile according to peer group adjusted alphas (APB). The joint application of AGT and APB methodologies eliminates the part of the winners selected following the stand-along models and reduces the number of the top winners to 546 funds.

3. Analysis of the Results

In this section, we test whether the combined AGT&APB procedure enables us to pick winner (loser) funds more accurately than the stand-along AGT or APB models. **We hypothesize that** 1) the investors selecting funds ranked at the top (bottom) based on the combined procedure should generate better (worse) performance one year ahead measured by average annualised Sharpe ratios, AGT and APB alphas (Table 3); 2) the top (bottom) quartile of funds formed using AGT *and* APB alpha jointly should generate statistically different performance to the top (bottom) quartile of funds formed using AGT or APB alpha separately (Table 3); and 3) the performance of funds in the top quartile should be significantly different from that in the bottom quartile of funds (Table 4).

3.1. Comparison of AGT, ABP and joint AGT&APB procedure

The results in Table 3 are presented in four panels, A-D, each corresponding to one of the four Morningstar categories reported in Table 1. Within each category, funds are ranked into quartiles by 36-monthly estimates of i) AGT alpha, ii) APB alpha and iii) jointly AGT and APB alpha. Each panel reports the average annualised Sharpe ratios, AGT alpha and APB alpha estimated one year post ranking for the top and the bottom quartile of funds. ***, ** and * represent 1%, 5% and 10% level of significance of the Sharpe ratios, AGT and APB alphas. Values in "bold" represent the cases when performance of funds ranked by AGT and APB model jointly is significantly different to the corresponding performance of funds ranked by AGT *or* APB model separately[9].

The results in Table 3 corroborate our expectation that a combination of models, i.e. ranking of funds in the top quartile according to AGT *and* APB alpha jointly will enable investors to pick funds with better future performance than if the ranking was done based on AGT or APB model separately. The table reveals that an investor who selects the top-performing funds based on their highest positive AGT *and* APB alphas will generate higher Sharpe ratios than those investors that carry out fund selection utilising only one of the two alpha measures only. For instance, within the Large-Cap Growth category in Panel A, the average annualized Sharpe ratio for the top quartile of funds with jointly positive AGT and APB alphas (0.9413) is higher than that calculated for the top quartile based on AGT alphas (0.7819) or APB alphas (0.8213)

[9] Detailed results of significance of these differences are available on request from authors.

separately. These Sharpe ratios are significant at the 1% level. The highest average Sharpe ratio is generated in the Large Cap Blend group.

In the bottom quartile, results show that combining AGT *and* APB models when ranking funds is a better way to identify future loser funds than using either of the models separately. The average one-year ahead Sharpe ratios in the bottom quartile formed by using AGT and APB alphas jointly are lower than comparable Sharpe ratios of bottom-quartile funds based on AGT and APB alphas separately. This result holds in all four panels of Table 3.

Each panel of Table 3 also reports the average annual AGT and APB alpha of top and bottom quartile funds one-year post-ranking. Note that AGT alpha one-year-ahead is estimated if the ranking was made based on AGT alpha or AGT and APB alpha jointly, while the APB alpha one-year-ahead is estimated when funds are ranked by APB alpha or both AGT and APB model jointly. Let us look at Panel A - the Large Cap Growth category where the relevant benchmark index (used in the AGT model) is Russell 1000 Growth. The average annualised AGT alpha one-year post-ranking is significantly negative (-2.479%, significant at 1%) when historical AGT alpha is used to rank funds into the top quartile; it becomes insignificant when funds are ranked jointly using the AGT and APB models. Similarly, selecting funds based on historical APB alpha will yield an investor 0.620% APB alpha one year ahead (significant at 10%), while selecting funds using AGT and APB alpha jointly will earn average annual alpha of 1.971% (significant at 1%) in the top quartile of funds. In the bottom quartile of Panel A, the joint model can identify the funds with more negative and/or more significant annual average AGT and APB alphas one-year-ahead. The same is found in all three remaining Panels of Table 3.

All the differences in Sharpe ratios, across the models (AGT&APB jointly vs AGT or APB separately) across Morningstar categories in Panels A-D, for both the top and bottom quartile, are significant at the 1% level[10]. The largest difference in the Sharpe ratios for top-quartile funds is found in Panel A, between the combined AGT and APB model ranking vs. ranking using AGT only (difference significant at 1%). For the bottom quartile, the largest difference in one-year-ahead Sharpe ratios is in Panel D (difference significant at 1%) between the funds ranked using AGT and APB jointly and those ranked using APB only. Similar can be said when AGT or APB alpha is used to gauge one-year-ahead performance, except for a somewhat

[10] Significance of the differences in Sharpe ratios, AGT or APB alphas is available on request.

weaker (insignificant) difference in one-year ahead APB alphas reported in Panel C, the Large Cap Value category and for top-performing funds in Panel D, Mid Cap category, albeit exhibiting the expected sign. The largest difference in alphas for top performing funds is reported for Mid-Cap Category (Panel D) between the funds ranked on AGT alphas vs. AGT and APB alphas jointly (1.622%, significant at 10%). Amongst the worst-performing funds, the largest difference (-1.632%, significant at 1%) is also recorded in Panel D, and it corresponds to the difference between the APB alpha of funds ranked solely by the APB model vs. those ranked by AGT and APB model jointly. This is not surprising as mid-cap stocks are relatively under researched asset class by analysts[11], with more opportunities to identify mispricing. Hence identifying 'true' mid-cap winners (losers) leads to stronger magnitude of improvement in performance.

---Insert Table 3 here---

Overall, Table 3 shows that our two-fold procedure of jointly using AGT and APB models identifies "clear" true winners and true losers, enabling us to select the funds for the top (bottom) quartile, that will generate more positive (negative) performance one year ahead compared to funds selected based on stand-alone AGT or ABP models. These findings are in contrast with studies where persistence is gauged on traditional factor models that find only underperformance persists (e.g. Carhart, 1997, Cuthbertson et al., 2008), but are in line with more recent literature utilising models employed in our two-fold fund selection process, such as Hunter et al. (2014) and Mateus et al. (2019b) which provide evidence of persistence of both winners and losers.

[11] Evidence from the industry supports this notion, see for instance : https://www.westernsouthern.com/touchstone/insights/the-unique-positioning-of-mid-cap-investing (Accessed: 22/10/2022)

3.2. Comparison of top and bottom quartile funds ranked by AGT only, ABP only and AGT and APB jointly

Our next step is to examine whether funds selected as winners and losers truly generate different performances. Table 4 reports the *differences* in Sharpe ratios, AGT alphas and APB alphas between the top and the bottom quartile of funds, using the ranking methods as per Table 3. All four panels of Table 4 show that AGT or ABP methods used in isolation are not very powerful in differentiating between the top and bottom performing funds one-year post-ranking. Ranking funds using the combined AGT&APB procedure creates a greater difference in the performance of the top and bottom quartile across all measures of performance one year ahead and in all four Morningstar categories (panels). More specifically, the difference in average Sharpe ratios one-year post-ranking between the top and bottom quartile of funds, when ranking is based on the AGT and APB models jointly, is statistically significant in all four panels. When ranking is based on either AGT or APB model separately, the difference in average one-year-ahead Sharpe ratios between the top and bottom quartile of funds is comparatively smaller and insignificant. The exception here is Panel D, the Mid-Cap category of funds, where the difference in Sharpe ratios between top and bottom quartile is significant for both AGT and APB ranking, but it is smaller than the difference in Sharpe ratios stemming from ranking based on both models jointly. Hunter et al. (2014) also document significant persistence in performance of Mid Cap funds in their APB approach, which is in line with the results presented.

Similar can be said when AGT alpha and APB alpha are used as a measure of performance one-year post-ranking. Overall, the strongest difference in using AGT and APB model simultaneously vs. using each model separately to rank the funds is observed in the Large Cap Growth category (Panel A), where all three measures of post-ranking performance show that ranking on AGT alpha only or APB alpha only will not lead to a significant difference between the funds ranked at the top and the bottom. In contrast, ranking them based on both AGT and APB alpha will enable investors to identify top-ranked funds that perform significantly better than bottom-ranked funds one year down the line. We are confident that look-ahead bias bares no significance to our results. First, the pattern of performance in our quartiles resembles an inverse of a J-shaped curve, as opposed to the U-shaped pattern documented by Horst et al.

(2001)[12] or the J-shaped pattern from Hendricks et al (1997)[13]. Second, our AGT and APB alphas are calculated one-year post ranking for funds with a minimum of 6 months of data so a small and approximately equal number of funds is eliminated in the top and bottom quartile post-ranking. Hence if look-ahead bias is present at all, it would drive performance down in both quartiles, not affecting their difference (i.e. top quartile will still have better performance). That is the case for all our ranking methods. Therefore, if any look-ahead bias were present, it would not impact the inferences regarding the two-fold AGT&APB procedure vis-a-vis the stand-along AGT or APB approaches.

---Insert Table 4 here---

3.3. False winners and false losers

In this section we look at the difference in one-year post-ranking performance between the top and bottom quartile of 'false winners (losers)' suggested by the stand-along AGT and APB models, however, eliminated by our two-stage combined approach. We claim that the difference should be positive if AGT or ABP rankings on their own are successful in isolating 'true' future winners (losers). Our results demonstrate the opposite. The difference in one-year post-ranking performance between the top and bottom quartile of 'false winners (losers)' is negative for all style categories. More importantly, the combined AGT&APB application trims close to 40% of funds in winner/loser quartiles, which represents a substantial reduction in comparison to the winners(losers) suggested by the AGT and APB models in isolation. Table 5 reports results for all style categories of funds. All differences in one-year post-ranking Sharpe ratios, AGT or APB alphas are negative, and some significant (around 60% of the cases). The greatest difference in annualized alphas is -5.37% (significant at 1% level), between the 'false' top and bottom quartile of Large Cap Value funds' AGT alphas (ranked on AGT only). The greatest difference in the Sharpe ratios is associated with the same category of funds, signalling that the AGT model ranking makes the most errors in the Large-Cap Value category.

---

[12] Where outperformance is found in both top and bottom quartile. Horst et al. (2001) claim that funds that take risks unaccounted for by Jensen's model are the ones that perform better if they survive 36 months post-ranking, creating a U-shaped pattern or performance across portfolio octiles. However, when applying Carhart four-factor model, the look-ahead bias in their study is found to be negligible.

[13] Where outperformance is found in the bottom quartile

Therefore, a separate application of AGT or APB models can result in the selection of false winner funds whose performance one year ahead is worse than that of the false loser funds. [14]

---Insert Table 5 here---

3.4. The economic value of performance one year ahead

While excess returns are not a risk-adjusted performance measure, they illustrate the economic value of performance. For instance, a fund that has the highest benchmark-adjusted alpha can still have a negative excess return; a scenario any investor would rather avoid. Table 6 summarizes results based on annualised excess returns. Panels A-D in Table 6 corresponds to four peer groups. Each panel shows in percentage the annual excess returns and the difference in annual excess returns across models (one-year post-ranking), and the difference in annual excess returns between the top and the bottom quartile. The results confirm the findings reported in Tables 3 and 4. For instance, in Panel A, the average excess return in the top quartile of Large Cap Growth funds following ranking based on AGT and APB jointly is 10.584% (significant at 5%); which is higher than the excess return of 8.095% (significant at 10%) following AGT ranking or 9.088% (significant at 5%) following APB ranking. The differences in excess returns across these models used for ranking of funds (AGT, APB or AGT&APB) are significant at a minimum 5% level with most differences being significant at a 1% level across panels A-D. Therefore, we show that when using both AGT and APB models, one can select top quartile funds that will generate the highest average annual excess return one year later. Moreover, only ranking funds using AGT and APB alpha jointly will generate statistically significantly different excess returns between the top and the bottom quartile funds one year ahead; using either AGT or APB model separately - will not. For example, looking at the Mid Cap Category of funds (Panel D), the difference in average annual excess return one year ahead between the top and bottom quartile of funds ranked using AGT and APB alpha jointly is 6.911% (significant at 1%), while that difference is smaller and statistically insignificant when either AGT (1.052%) or APB (2.290%) is used to sort the funds into quartiles. The same is observed in all other panels.

[14] The list of false winners/losers is available upon request.

---Insert Table 6 here---

4. Robustness tests

4.1. Performance three years post-ranking

In the first of our robustness tests, we extend our post-ranking window of measuring the performance of the top and bottom-ranked funds to 36 months to match a pre-estimation window and test longer-term persistence, as longer time series are expected to result in more accurate alpha estimates (e.g. Huij and Veerbek, 2007). To do this, we select all the funds that have a minimum of 30 months of data within a 36-month post-ranking window. Such data requirement reduces the number of funds across all categories, for instance, the number of funds in the large-cap blend category changes from 841 to 733, and the number of rolling periods decreases from 21 to 19. Next, we rank the funds into quartiles each year. The findings obtained are qualitatively similar to the ones from Table 3. They confirm our previous results indicating that the application of both AGT&APB models acts as a better predictor of future winners than each of the models on their own. Notably, alphas of the top quartile funds resulting from the combined AGT&APB ranking are all significantly higher in the top than in the bottom quartile, which is not the case when winners are selected based on AGT or APB only. The differences in Sharpe ratio, AGT and APB alphas in the top quartile[15] are significant across ranking procedures used. The results are consistent across all style categories and available upon request.

4.2. Fixed effects panel estimation of post-ranking performance

Our alphas in Tables 3-6 are all mean values of fund-by-fund regressions in post-ranking periods of either t+12m or t+36m. To confirm that those average values reported are not affected by distribution in the quartiles, we estimate one-year-ahead (t+12) and three-year ahead (t+36) alphas using a fixed effect panel model (Mateus et. al., 2016). We repeat this for funds ranked based on AGT, APB, and a combination of both. Our findings are consistent with the previous results. Funds selected based on the combined application of models generate positive and significant (at 10%) alpha in the following year. Similarly, one-year-ahead alphas stemming from AGT&APB ranking are higher (for top quartile) or lower (for bottom quartile)

[15] The significance of the differences is less pronounced in the bottom quartile.

vs. AGT or APB in isolation. The differences between top/bottom quartiles from combined models' ranking are by and large positive and statistically significant. Findings for all categories of funds are qualitatively the same and are available on request.[16]

4.3. Decile portfolios

As a final robustness test, we repeat the analysis using decile portfolios. The method enables us to isolate the most extreme winners (losers) in each style category. Due to our minimum requirement of 10 funds per decile each year, the top (bottom) 10% of funds selected in each style category are combined into a single top (bottom) decile. We refer to these deciles as Combined Categories Deciles[17], which are constructed for every year in our sample (21 in total). The results strongly support our findings from corresponding tables for quartile funds (Tables 3 and 4) and confirm the superiority of the joint ranking procedure for the selection of future winners.

4.4. Additional tests

In addition to all robustness tests described in sections 4.1.- 4.3., we replicate the results from Table 3 excluding 1, 2 and 3 rolling windows at a time. By doing this, we test whether any particular period, such as the dot.com boom or the 2008-2010 financial crisis bares a strong influence on our results. We find that our conclusions are not driven by any particular period in our sample[18]. These results are available from authors on request.

In conclusion, all the robustness tests in this section corroborate our main inferences from Section 3 and confirm that they are not sensitive to the post-ranking estimation period, not driven by outliers or by the size of fund sorts.

---

[16] We also include year dummies and results are qualitatively similar. Tables are available upon request.

[17] Please note that if we have simply selected top (bottom) 10% of all funds each year to form our top (bottom) deciles, some of our winner (loser) deciles could be populated by a single style.

[18] In addition, results are not sensitive to the impact of outliers: we winsorize the data in each quartile, each year, and each style by trimming the outliers using a 5% and 95% cut-off point and find the same results from the repeat analysis.

5. Conclusions

The literature on mutual fund performance and performance persistence overwhelmingly support the fact that persistence, particularly among winner funds, is rare. Academic evidence also suggests that winner funds are those that generate the highest three/four-factor alphas or some other performance measure (such as Sharpe ratio for instance). However, the performance of fund managers in the industry is gauged against a fund's self-reported prospectus benchmark and the peer-group. Hence, in this paper, we propose a new two-fold procedure that allows identifying "true" winner and loser funds: those that outperform the benchmark *and* are at the top of their peer group at the same time, based on factor-risk-adjusted measures of performance. Our method differs from existing studies that either gouge benchmark-adjusted performance only (e.g. Angelidis et al., 2013) or peer-group adjusted performance only (e.g. Hunter et al., 2014). We contribute to the literature on mutual fund performance and persistence in performance by demonstrating that using the two novel approaches hand-in-hand, we allow investors to eliminate the 'false' winners (losers), suggested by one but not both of the models - thus focusing on 'true' winners (losers) that outbid (fail) both the benchmark and the peer group.

We illustrate on the sample of 989 US active equity mutual funds grouped into peer groups according to four Morningstar Global Categories: Large Cap Growth, Large Cap Blend, Large Cap Value and Mid Cap. For each fund, we estimate benchmark-adjusted and peer-group-adjusted alphas following Angelidis et al. (2013) approach (AGT) and Hunter et al. (2014) method (APB) respectively. One could argue that a limitation of this study could be the choice of funds' benchmark and the choice of the peer-group. We use well established Morningstar fund categorisation based on portfolio holdings to determine the peer group for our funds and well-known US benchmark indices widely used as proxies for the above investment styles as funds' benchmarks. We strongly believe that this approach by and large mitigates the benchmark or peer-group specification problem.

In our method, we rank the funds within each Morningstar global category, and form quartiles at the start of every year according to funds' historical 1) AGT alpha, 2) APB alpha and 3) using our two-step approach in which we utilise both AGT and APB alphas. We test the performance of the top and bottom quartile funds one-year post-ranking using average annualised Sharpe ratios, AGT alphas and APB alphas. Our findings confirm the superiority of our two-fold procedure in several ways. First, across all Morningstar categories, higher Sharpe ratios, AGT alphas, APB alphas and excess returns are achieved when top funds are selected based on AGT and APB alphas jointly. Second, when both models are utilized for fund ranking, there is a significant difference in one-year-ahead performance between top and bottom-ranked funds. The distinction is less pronounced when stand-along AGT and APB models are used. Further, we document that 40 percent of funds selected based on the stand-along AGT and APB models are “false” winners and losers. Those funds generate worse (better) performance than the funds in the top (bottom) quartile and get eliminated when our joint AGT&APB selection procedure is used.

We also consider the economic value of performance by estimating annual excess returns of funds and confirm our main findings. Our results are robust to a variety of tests: the extended post-ranking investment horizon of 36 months; the alpha estimation method; they are not driven by outliers in each quartile, the size of fund-sorts when quartiles are replaced by deciles; and, finally, they are not sensitive to any particular period within our sample.

This paper contributes to the academic literature on active equity fund performance persistence by providing evidence that our suggested two-stage combined procedure allows investors to identify ‘true’ winners with consecutive performance persistence up to three years ahead. Our results supporting persistence in performance of ‘true’ winners and losers are in line with Mateus et al. (2019a) and Hunter et al. (2014) who utilise only peer-group adjusted model, but we demonstrate that adding benchmark-adjusted model in the fund selection process further improves persistence. Our findings aim to raise awareness among investment practitioners and researchers regarding the implications of the choice of mutual fund ranking models on subsequent persistence in performance. Our approach can be extended to any country, market and peer-group of funds, such as fixed income, sector-specific, smart-beta ETFs etc.

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., 2015. Investment Style Misclassification and Mutual Fund Performance, SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2648375

Brown S. J. and Goetzmann W. N., 1995. Performance Persistence, *Journal of Finance*, 50 (2), pp. 679-698.

Carhart, M. M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, pp. 57-82.

Chinthalapati, V.L., Mateus, C. and Todorovic, N., 2017, Alphas in disguise: A new approach to uncovering them, *International Journal of Finance and Economics*, 22 (3), pp. 234-243.

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

Cremers, K. J. M., Fulkerson, J. A. and Riley, T. B., 2019, Challenging the Conventional Wisdom on Active Management: A Review of the Past 20 Years of Academic Literature on Actively Managed Mutual Funds, *Financial Analysts Journal*, 75(4), pp. 8-32.

Cremers, K.M., Fulkerson, J.A. and Riley, T.B., 2022. Benchmark discrepancies and mutual fund performance evaluation. *Journal of Financial and Quantitative Analysis*, *57*(2), pp.543-571.

Cochrane, J.H., 2011. Presidential address: Discount rates. The Journal of finance, 66(4), pp.1047-1108.

Cuthbertson, K., Nitzsche D. and O'Sullivan N., 2008, UKMutual Fund Performance: Skill or Luck?, *Journal of Empirical Finance* 15(4), pp. 613-634.

Cuthbertson, K., Nitzsche D. and O'Sullivan N., 2010, Mutual fund performance: measurement and evidence, *Financial Markets, Institutions and Instruments*, 19 (2), pp. 95-187.

Davies, J. L. (2001), Mutual Fund Performance and Manager Style), *Financial Analysts Journal*, Vol. 57, No. 1 (Jan. - Feb., 2001), pp. 19-27

Ding, H., Zheng, H., Zhu, Ch. (2015) Equity funds in emerging Asia: Does size matter? International *Review of Economics & Finance*, 35(3), 149–165.

Elton, E., Gruber, M., Blake, C.,1996, The Persistence of Risk-Adjusted Mutual Fund Performance, *The Journal of Business,* 69(2), pp.133-157.

Elton, E., Gruber, M., Blake, C., 2003. Incentive fees and mutual funds. *Journal of Finance* 58, pp.779–804

Elton, E., Gruber, M. 2011. Mutual funds. SSRN *working paper*: https://ssrn.com/abstract=2088418

Elton, E., Gruber, M. and de Souza A., 2019, Are Passive Funds Really Superior Investments: An Investor Perspective, *Financial Analysts Journal*, 75(3), pp.7-19.

Elton, E., Gruber, M. 2020, A Review of the Performance Measurement of Long-Term Mutual Funds, *Financial Analysts Journal*, 76(3), pp. 22-37.

Fama, E. F., and French, K. R., 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, pp. 3-56.

Fama, E.F. and French, K.R., 2010, "Luck Versus Skill in the Cross-Section of Mutual Fund Returns." *Journal of Finance*, 65(5): pp.1915-47

Ferreira, M.A., Keswani, A., Miguel, A.F., Ramos, S. (2013) The determinants of mutual fund performance: A cross-country study, *Review of Finance*, 17(2), 483–525.

Filip, D., Rogala, T. (2020) Analysis of Polish mutual funds performance: a Markovian approach, *Statistics in Transition new series*, 22(1), 115-130.

Fletcher, J. and D. Forbes, 2002. "An Exploration of the Persistence of UK Unit Trusts Performance. *Journal of Empirical Finance*, 9, pp. 475-493.

Foran, J. and O'Sullivan, N., 2014. Liquidity risk and the performance of UK mutual funds. *International Review of Financial Analysis*, 35, pp.178-189

Frazzini, A., Friedman, J. and Pomorski, L., 2016. Deactivating active share. *Financial Analysts Journal*, 72(2), pp.14-21.

French, Kenneth R., 2008. The Cost of Active Investing, *Journal of Finance* 63, 1537–1573

Gallagher, D.R. (2003) Investment manager characteristics, strategy, top management changes and fund performance, *Accounting and Finance*, 43(3), 283–309.

Hunter, D., Kandel E., Kandel S. and Wermers R., 2014, Mutual fund performance evaluation with active peer benchmarks, *Journal of Financial Economics*, 112(1), pp. 1-29.

Horst, J.R., Nijman,T.E. and Verbeek M., 2001, Eliminating look-ahead bias in evaluating persistence in mutual fund performance, *Journal of Empirical Finance*, 8(2001), pp.345-373.

Hendricks, D., Patel, J., Zeckhauser, R., 1997. The J-shape of performance Persistence given Survivorship Bias. *The Review of Economics and Statistics* 79, 161–170.

Jiang W., 2003. A nonparametric test of market timing, Journal of Empirical Finance**,** Volume 10, Issue 4, September 2003, Pages 399–425

Matallin-Saez, J.C. (2007), Portfolio Performance: Factors or Benchmarks ? *Applied Financial Economics*, 17:14, pp.1167–1178.

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.

Mateus, I.B., Mateus, C. and Todorovic, N. A (2019a). Benchmark-adjusted performance of US equity mutual funds and the issue of prospectus benchmarks. Journal of Asset Management, 20(1), pp. 15–30

Mateus, I.B., Mateus, C. and Todorovic, N., (2019b). Use of active peer benchmarks in assessing UK mutual fund performance and performance persistence. *The European Journal of Finance*, *25*(12), pp.1077-1098.

Mateus, I.B., Mateus, C. and Todorovic, N. (2019c). Review of new trends in the literature on factor models and mutual fund performance. International Review of Financial Analysis, 63, pp. 344–354.

Moreno D, Rodríguez R., 2009. The value of coskewness in mutual fund performance evaluation Journal of Banking & Finance, September, Pages 1664–1676

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.

Wermers R., 2000. Mutual Fund Performance: An Empirical Decomposition into Stock-Picking Talent, Style, Transactions Costs, and Expenses, The Journal of Finance, 55 (4): 1655–1703.

Table 1: Final Sample

Benchmark	**Morningstar Category**	**# Funds**	**# Monthly Observations**
S&P 500	US Large Cap Blend	460	73,493
RUSSELL 1000 GROWTH	US Large Cap Growth	290	48,393
RUSSELL 1000 VALUE	US Large Cap Value	127	21,160
RUSSELL MIDCAP	US Mid Cap	112	17,332
		989	**160,378**

Table 2: **Top and Bottom quartile funds with the available one-year-post-ranking data, stand-along AGT and APB application and the combined procedure**

The table shows the number of funds that rank in the top and bottom quartile according to i) AGT alpha ii) APB alpha and iii) both AGT and APB alphas estimated using historical 36-months data. The number reflects the aggregate number of funds in 21 rolling periods with minimum 6 months returns in the year ahead.

Peer group / Benchmark	***Top Quartile***			***Bottom Quartile***		
	AGT model	***APB model***	***AGT + APB model***	***AGT model***	***APB model***	***AGT + APB model***
Large Cap Growth/ Russell 1000 Growth	695	717	546	696	713	563
Large Cap Blend /S&P 500 Index	1114	1120	841	1112	1114	823
Large Cap Value/ Russell 1000 Value	326	325	250	327	321	251
Mid Cap / Russell Mid Cap	235	239	175	231	232	172

Table 3: One-year-ahead Sharpe ratios, AGT alphas and APB alphas of Top and Bottom quartile of funds, by model and peer group/benchmark

Each panel of the table shows one Morningstar category and corresponding benchmark: Large Cap Growth (Russell 1000 growth), Large Cap Blend (S&P 500), large Cap Value (Russell 1000 Value) and Mid Cap (Russell Mid Cap). Funds are ranked into quartiles by 36-monthly estimates of i) AGT alpha, ii) APB alpha and iii) jointly AGT and APB alpha. Within each panel we report the average Sharpe ratios, AGT alpha and APB alpha estimated one year post ranking for the top and the bottom quartile of funds. Note that AGT (APB) alpha one year ahead is estimated only if funds are ranked into quartiles by historical AGT (APB) alphas or by AGT and APB alpha jointly. t-statistics are in parentheses, where ***, ** and* correspond to 1%, 5% and 10% level of significance of the Sharpe ratios, AGT and APB alphas. Values in **bold** represent the cases when performance of funds ranked by AGT and APB model jointly is significantly different to the corresponding performance of funds ranked by AGT *or* APB model separately. Detailed results of significance of these differences are available on request.

Panel A: Large Cap Growth, Russell 1000 Growth Index					
Top Quartile			*Bottom Quartile*		
Annual Average Sharpe Ratio one year post ranking, by model			*Annual Average Sharpe Ratio* one year post ranking, by model		
AGT model	*APB model*	*AGT+APB model*	*AGT*	*APB*	*AGT+APB model*
*0.7819**** *(2.89)*	*0.8213**** *(2.95)*	***0.9413***** ***(3.33)***	*0.8271**** *(2.94)*	*0.8171**** *(2.92)*	***0.7439***** ***(2.67)***
Annual AGT alpha, one-year post ranking*, by model*			*Annual AGT alpha, one-year* post ranking*, by model*		
AGT model *-2.479%**** *(-3.47)*		*AGT +APB model* ***-0.998%*** ***(-1.30)***	*AGT model* *-2.725%**** *(-4.35)*		*AGT +APB model* ***-3.493%****** ***(-5.12)***
Annual APB alpha, one-year post ranking*, by model*			*Annual APB alpha, one-year* post ranking*, by model*		
APB model *0.620%** *(2.00)*		*AGT +APB model* ***1.971%****** ***(4.22)***	*APB model* *-0.296%* *(-0.69)*		*AGT +APB model* ***-0.957%***** ***(-2.28)***

Panel B: Large Cap Blend, S&P 500 Index					
Top Quartile			*Bottom Quartile*		
Annual Average Sharpe Ratio one year post ranking, by model			*Annual Average Sharpe Ratio* one year post ranking, by model		
AGT model	*APB model*	*AGT+APB model*	*AGT model*	*APB model*	*AGT+APB model*
*0.8671**** *(3.01)*	*0.8533**** *(2.99)*	***0.9777***** ***(3.34)***	*0.8237**** *(2.89)*	*0.8322**** *(2.91)*	***0.7813***** ***(2.76)***
Annual AGT alpha, one-year post ranking*, by model*			*Annual AGT alpha, one-year* post ranking*, by model*		
AGT model *-1.305%**** *(-3.29)*		*AGT +APB model* ***-0.521%*** ***(-1.15)***	*AGT model* *-2.347%**** *(6.48)*		*AGT +APB model* ***-2.619%****** ***(-7.08)***
Annual APB alpha, one-year post ranking*, by model*			*Annual APB alpha, one-year* post ranking*, by model*		
APB model *0.234%* *(0.99)*		*AGT +APB model* ***1.008%***** ***(2.65)***	*APB model* *-0.610%**** *(-3.58)*		*AGT +APB model* ***-1.190%****** ***(-5.22)***

Panel C: Large Cap Value, Russell 1000 Value Index

Top Quartile			Bottom Quartile		
Annual Average Sharpe Ratio one year post ranking, by model			*Annual Average Sharpe Ratio* one year post ranking, by model		
AGT model	*APB model*	*AGT+APB model*	*AGT model*	*APB model*	*AGT+APB model*
*0.7180*** *(2.80)*	*0.7193*** *(2.83)*	***0.8120****** ***(3.13)***	*0.6896*** *(2.65)*	*0.6921*** *(2.66)*	***0.6131**** ***(2.35)***
Annual AGT alpha, one year post ranking, *by model*			*Annual AGT alpha, one year* post ranking, *by model*		
AGT model *-0.320%* *(-0.42)*		*AGT +APB model* ***0.430%*** ***(0.56)***	*AGT model* *-0.471%* *(-0.68)*		*AGT +APB model* ***-0.881%*** ***(-1.10)***
Annual APB alpha, one year post ranking, *by model*			*Annual APB alpha, one1 year* post ranking, *by model*		
APB model *0.016%* *(0.04)*		*AGT +APB model* *0.421%* *(0.93)*	*APB model* *-0.579%* *(-0.89)*		*AGT +APB model* *-0.703%* *(-0.8)*

Panel D: Mid Cap, Russell Mid Cap					
Top Quartile			*Bottom Quartile*		
Annual Average Sharpe Ratio one year post ranking, by model			*Annual Average Sharpe Ratio* one year post ranking, by model		
AGT model	*APB model*	*AGT+APB model*	*AGT model*	*APB model*	*AGT+APB model*
*0.7110*** *(2.19)*	*0.7206*** *(2.19)*	***0.8364**** ***(2.51)***	*0.4989* *(1.70)*	*0.5069* *(1.72)*	***0.3928*** ***(1.39)***
AGT alpha, one year post ranking, *by model*			*AGT alpha, one year* post ranking, *by model*		
AGT model *-2.242%*** *(-2.59)*		*AGT +APB model* ***-0.620%*** ***(-0.46)***	*AGT model* *-5.133%**** *(-5.12)*		*AGT +APB model* ***-6.725%****** ***(-7.09)***
APB alpha, one year post ranking, *by model*			*APB alpha, one year* post ranking, *by model*		
APB model *0.503%* *(1.08)*		*AGT +APB model* *1.519%*** *(2.55)*	*APB model* *-1.754%**** *(-3.43)*		*AGT +APB model* ***-3.386%****** ***(5.22)***

Table 4: Differences in Sharpe ratios, AGT and APB alphas between Top and Bottom quartile of funds, by ranking model

Each panel of the table shows one Morningstar category and corresponding benchmark: Large Cap Growth (Russell 1000 growth), Large Cap Blend (S&P 500), large Cap Value (Russell 1000 Value) and Mid Cap (Russell Mid Cap). Funds are ranked into top and bottom quartile in year t using their historical 36-monthly i) AGT alpha, ii) APB alpha and iii) AGT and APB alpha jointly. Sharpe ratios, AGT alphas and APB alphas are estimated for top and bottom quartile one-year post ranking. This table shows the difference in Sharpe ratios, AGT and APB alphas between the top and bottom quartile of funds, by each ranking model used. Note that AGT (APB) alpha one-year-ahead is estimated only if funds are ranked into quartiles by historical AGT (APB) alphas or by AGT and APB alpha jointly. T-statistics are in parentheses. ***, ** and* are corresponding to 1%, 5% and 10% level of significance.

Panel A: Large Cap Growth, Russell 1000 Growth Index		
Top vs. Bottom Quartile difference in Sharpe Ratios 1-year post ranking, by model		
AGT	***APB***	***AGT+APB model***
-0.0453 *(-1.09)*	*0.0042* *(0.12)*	***0.1974****** ***(5.06)***
Top vs. Bottom Quartile difference in AGT alphas 1-year post ranking, by model		
AGT model		***AGT+APB model***
0.245% *(1.52)*		***2.495%****** ***(4.76)***
Top vs. Bottom Quartile difference in APB alphas 1-year post ranking, by model		
APB model		***AGT+APB model***
0.916% *(1.52)*		***2.928%****** ***(4.19)***

Panel B: Large Cap Blend, S&P 500 Index		
Top vs. Bottom Quartile difference in Sharpe Ratios 1-year post ranking, by model		
AGT	***APB***	***AGT+APB***
0.0434 *(1.17)*	*0.0211* *(0.57)*	***0.1964****** ***(5.40)***
Top vs. Bottom Quartile difference in AGT alphas 1-year post ranking, by model		
AGT model		***AGT+APB model***
*1.043%**** *(3.21)*		***2.097%****** ***(4.87)***
Top vs. Bottom Quartile difference in APB alphas 1-year post ranking, by model		
APB model		***AGT+APB model***
*0.845%*** *(2.51)*		***2.198%****** ***(4.51)***

<table>
<tr><td colspan="4">Panel C: Large Cap Value, Russell 1000 Value Index</td></tr>
<tr><td colspan="4">Top vs. Bottom Quartile difference in Sharpe Ratios 1-year post ranking, by model</td></tr>
<tr><td>AGT</td><td>APB</td><td colspan="2">AGT+APB</td></tr>
<tr><td>0.0284
(0.84)</td><td>0.0272
(0.60)</td><td colspan="2">0.1988***
(5.45)</td></tr>
<tr><td colspan="4">Top vs. Bottom Quartile difference in AGT alphas 1-year post ranking, by model</td></tr>
<tr><td colspan="2">AGT model</td><td colspan="2">AGT+APB model</td></tr>
<tr><td colspan="2">0.151%
(0.28)</td><td colspan="2">1.311%*
(2.01)</td></tr>
<tr><td colspan="4">Top vs. Bottom Quartile difference in APB alphas 1-year post ranking, by model</td></tr>
<tr><td colspan="2">APB model</td><td colspan="2">AGT+APB model</td></tr>
<tr><td colspan="2">0.595%
(0.75)</td><td colspan="2">1.124%
(1.32)</td></tr>
</table>

<table>
<tr><td colspan="4">Panel D: Mid Cap, Russell Mid Cap</td></tr>
<tr><td colspan="4">Top vs. Bottom Quartile difference in Sharpe Ratios 1 year post ranking, by model</td></tr>
<tr><td>AGT</td><td>APB</td><td colspan="2">AGT+APB</td></tr>
<tr><td>0.2122***
(3.08)</td><td>0.2137**
(2.56)</td><td colspan="2">0.4437***
(5.16)</td></tr>
<tr><td colspan="4">Top vs. Bottom Quartile difference in AGT alphas 1-year post ranking, by model</td></tr>
<tr><td colspan="2">AGT model</td><td colspan="2">AGT+APB model</td></tr>
<tr><td colspan="2">2.891%***
(3.90)</td><td colspan="2">6.105%***
(4.27)</td></tr>
<tr><td colspan="4">Top vs. Bottom Quartile difference in APB alphas 1-year post ranking, by model</td></tr>
<tr><td colspan="2">APB model</td><td colspan="2">AGT+APB model</td></tr>
<tr><td colspan="2">2.257%***
(2.82)</td><td colspan="2">4.905%***
(5.24)</td></tr>
</table>

Table 5: False Winners vs False Losers: differences in Sharpe ratios, AGT and APB alphas between top and bottom quartile of funds, one-year post ranking

In this table 'false winners (losers)' are funds that rank in the top (bottom) quartile according to AGT (or APB) model but not according to APB (or AGT) model. Table reports differences in AGT alphas if the funds are ranked top/bottom by AGT model (and not APB) and the differences in APB alphas when they are ranked top/bottom by APB model (but not AGT). Results are reported for all four style categories in this paper: Large Cap Growth, Large Cap Blend, Large Cap Value, and Mid Cap. T-statistics are in parentheses. ***, ** and* are corresponding to 1%, 5% and 10% level of significance.

Top vs. Bottom Quartile difference in Sharpe ratios1 year post ranking, by model			
Large Cap Growth, Russell 1000 Growth Index		***Large Cap Blend, S&P 500 Index***	
AGT model	***APB model***	***AGT model***	***APB model***
*-0.21016***	*-0.0531*	*-0.3688***	*-0.4517*
(-2.35)	*(-0.16)*	*(-2.13)*	*(-2.96)*
Large Cap Value, Russell 1000 Value Index		***Mid Cap, Russell Mid Cap***	
AGT model	***APB model***	***AGT model***	***APB model***
*-0.7982****	*-0.5771****	*-0.3808**	*-0.3364***
(-4.28)	*(-4.40)*	*(-2.03)*	*(-2.01)*
Top vs. Bottom Quartile difference in AGT/APB alphas 1 year post ranking, by model			
Large Cap Growth, Russell 1000 Growth Index		***Large Cap Blend, S&P 500 Index***	
AGT model	***APB model***	***AGT model***	***APB model***
-2.19706%	*-2.08589*	*-1.60165%*	*-2.33871***
(-1.39)	*(-1.36)*	*(1.72)*	*(-2.35)*
Large Cap Value, Russell 1000 Value Index		***Mid Cap, Russell Mid Cap***	
AGT model	***APB model***	***AGT model***	***APB model***
*-5.36895****	*-2.32771*	*-5.08695***	*-5.08752**
(-3.29)	*(-1.40)*	*(-2.23)*	*(-1.93)*

Table 6: Economic value of performance

Each panel of this Table shows the annual excess returns (in %) one year post-ranking, the difference in annual excess returns (in %) across models one year post-ranking and the difference in annual excess returns (in %) between the top and the bottom quartile. Panels A-D correspond to four peer groups: Large Cap Growth, Large Cap Blend, Large Cap Value and Mid Cap. T-statistics are in parentheses, where ***, ** and* are corresponding to 1%, 5% and 10% level of significance.

Panel A: Large Cap Growth, Russell 1000 Growth Index						
	Top Quartile			*Bottom Quartile*		
	***Excess Return* one year post ranking, by model**			***Excess Return* one year post ranking, by model**		
	AGT	***APB***	***AGT+APB***	***AGT***	***APB***	***AGT+APB***
Annual Average Excess Return (p.a. in %)	*8.095** *(1.99)*	*9.088*** *(2.15)*	*10.584*** *(2.46)*	*9.193*** *(2.21)*	*8.616*** *(2.09)*	*7.649** *(1.89)*
	Difference in Excess Return one year ahead across models			**Difference in Excess Return one year ahead across models**		
	AGT+APB vs. AGT	***AGT+APB vs. APB***		***AGT+APB vs. AGT***	***AGT+APB vs. APB***	
Difference in annual Excess Return (p.a. in %)	*2.489*** *(2.39)*	*1.497**** *(3.73)*		*-1.544**** *(-5.25)*	*-0.968**** *(-4.69)*	
	Top vs. Bottom Quartile difference in Excess Return, by model					
	AGT	***APB***	***AGT+APB***			
Difference in annual Excess Return (p.a. in %), Top vs. Bottom Quartile	*-1.098* *(-1.42)*	*0.471* *(0.89)*	*2.936**** *(3.87)*			

Panel B: Large Cap Blend, S&P 500 Index						
	Top Quartile			*Bottom Quartile*		
	***Excess Return* one year ahead by model**			***Excess Return* one year ahead by model**		
	AGT	***APB***	***AGT+APB***	***AGT***	***APB***	***AGT+APB***
Annual Average Excess Return (p.a. in %)	*7.907*** *(2.19)*	*7.844*** *(2.15)*	*9.184*** *(2.54)*	*7.806*** *(2.10)*	*7.929*** *(2.15)*	*7.329** *(1.98)*
	Difference in Excess Return one year ahead across models			**Difference in Excess Return one year ahead across models**		
	AGT+APB vs. AGT	***AGT+APB vs. APB***		***AGT+APB vs. AGT***	***AGT+APB vs. APB***	
Difference in annual Excess Return (p.a. in %)	*1.277**** *(9.89)*	*1.340**** *(5.98)*		*-0.478*** *(-2.65)*	*-0.601**** *(-4.29)*	
	Top vs. Bottom Quartile difference in Excess Return, by model					
	AGT	***APB***	***AGT+APB***			
Difference in annual Excess Return (p.a. in %), Top vs. Bottom Quartile	*0.101* *(0.21)*	*-0.086* *(-0.18)*	*1.855**** *(4.67)*			

Panel C: Large Cap Value, Russell 1000 Value Index						
	Top Quartile			***Bottom Quartile***		
	***Excess Return* one year ahead by model**			***Excess Return* one year ahead by model**		
	AGT	***APB***	***AGT+APB***	***AGT***	***APB***	***AGT+APB***
Annual Average Excess Return (p.a. in %)	*6.513** *(1.93)*	*6.606** *(1.99)*	*7.873*** *(2.34)*	*6.985** *(1.93)*	*6.970** *(1.96)*	*5.868** *(1.71)*
	Difference in Excess Return one year ahead across models			**Difference in Excess Return one year ahead across models**		
	AGT+APB vs. AGT	***AGT+APB vs. APB***		***AGT+APB vs. AGT***	***AGT+APB vs. APB***	
Difference in annual Excess Return (p.a. in %)	*1.359**** *(8.13)*	*1.267**** *(5.17)*		*-1.117**** *(-3.52)*	*-1.102*** *(-2.73)*	
	Top vs. Bottom Quartile difference in Excess Return, by model					
	AGT	***APB***	***AGT+APB***			
Difference in annual Excess Return (p.a. in %), Top vs. Bottom Quartile	*-0.471* *(-0.71)*	*-0.364* *(-0.46)*	*2.005**** *(3.19)*			

Panel D: Mid Cap, Russell Mid Cap						
	Top Quartile			***Bottom Quartile***		
	***Excess Return* one year ahead by model**			***Excess Return* one year ahead by model**		
	AGT	***APB***	***AGT+APB***	***AGT***	***APB***	***AGT+APB***
Annual Average Excess Return (p.a. in %)	*7.749* *(1.32)*	*8.826* *(1.49)*	*11.115** *(1.95)*	*6.697* *(1.32)*	*6.133* *(1.22)*	*4.204* *(0.85)*
	Difference in Excess Return one year ahead across models			**Difference in Excess Return one year ahead across models**		
	AGT+APB vs. AGT	***AGT+APB vs. APB***		***AGT+APB vs. AGT***	***AGT+APB vs. APB***	
Difference in annual Excess Return (p.a. in %)	*3.367**** *(3.90)*	*2.290*** *(2.76)*		*-2.493**** *(-3.36)*	*-1.929**** *(-5.10)*	
	Top vs. Bottom Quartile difference in Excess Return, by model					
	AGT	***APB***	***AGT+APB***			
Difference in annual Excess Return (p.a. in %), Top vs. Bottom Quartile	*1.052* *(0.60)*	*2.693* *(1.22)*	*6.911**** *(3.80)*			