Journal of Asset Management (2023) 24:284–298
https://doi.org/10.1057/s41260-023-00310-7

ORIGINAL ARTICLE

UK mutual funds: performance persistence and portfolio size

Keith Cuthbertson[1] · Dirk Nitzsche[1] · Niall O'Sullivan[2]

Revised: 23 January 2023 / Accepted: 4 April 2023 / Published online: 4 May 2023

Abstract
We re-examine performance persistence amongst UK mutual funds. Specifically, we investigate performance persistence amongst small portfolios of past high-performing funds. In contrast to the more common analysis of decile portfolios of funds, we focus on persistence in the more extreme positive tail of the cross section of fund performance. This paper contributes to the smaller literature on UK rather than US mutual fund performance. We investigate fund persistence based on practitioner index models as well as academic factor models, focusing on small portfolios of funds using inference based on nonparametric persistence test statistics as well as conventional t tests. We provide strong evidence of positive persistence amongst small-size portfolios of (past) high-performing funds that is robust to alternative formation and holding periods and alternative performance models. We also document some sensitivity in inferences on positive persistence when using nonparametric versus conventional tests.

Keywords Mutual fund performance persistence · Factor models · Portfolio size

JEL Classification G11 · G12 · G14 · C15

Introduction

In this paper, we re-examine UK mutual fund performance persistence (1990–2017) in the context of alternative factor models, concentrating on small portfolios of funds in the extreme tail of the cross-sectional distribution of fund performance (e.g. top-performing portfolios of size 2,3,5,7,10,20,35,50 mutual funds). We conjecture that there may be stronger positive persistence in small-size portfolios relative to the more commonly studied larger decile portfolios.

A number of studies have highlighted the need to consider fund performance persistence using both "academic factor" models (e.g. Fama and French 2015, 2016, 2017) and "practitioner index" models (e.g. the 4-factor IDX4 and 7-factor IDX7 index-based models of Cremers et al. 2013). Unlike the Fama–French academic factors, these index models generally have zero alphas with respect to a wide range of passive indices and style-sorted stock portfolios. Academic factor models use "risk factors" which have explanatory power for the cross section of average stock returns, whereas practitioner factors are chosen to more closely represent low-cost investible passive indices (which do not necessarily price stock portfolios).

Non-normality in fund returns (and factor model residuals) can arise in the tails of the cross-sectional distribution of funds, particularly for small-size portfolios formed on the basis of short backward-looking formation periods. We therefore implement a nonparametric bootstrap procedure that adjusts for non-normality and contrast these bootstrap statistical inferences with those from using conventional t tests. We consider fund performance persistence based on both net-of-fee fund returns (net alphas) and gross returns (gross alphas,) thus separating manager performance from investor abnormal performance persistence.

We examine the robustness of mutual fund performance persistence over alternative portfolio formation and holding periods (f, h). Past "winner" portfolios are sorted on past t-alphas for $f = 36$ or 60 months and h= 1,3,6 and 12 months, either into decile-size portfolios or into portfolios of a small number of funds s_i, $(i = 2, 3, 5, 7, 10, 20, 35, 50)$.

✉ Niall O'Sullivan
niall.osullivan@ucc.ie

1 Bayes Business School, City, University of London, London, UK

2 Cork University Business School, University College Cork, Cork, Ireland

There has not been an extensive recent UK study assessing the robustness of positive persistence in mutual fund portfolios, with respect to small versus large portfolios, different formation and holding periods, different factor models and with respect to standard t test versus bootstrap t values.

Our key result is that UK funds sorted on past t-alphas result in positive net-alpha persistence for small portfolios of past winner funds. In addition we find larger post-sort positive net alphas for formation periods based on 60 rather than 36 observations and for relatively short holding periods of up to 6 months. When using our bootstrap procedure, these inferences are robust to non-normality in funds' idiosyncratic risks. Hence, a fairly simple investment strategy of sorting UK funds into small portfolios based on past performance delivers positive net alphas for investors. This can be contrasted with previous studies that use many alternative sorting rules, including multiple sorts, which may be prone to data mining and false discoveries.

Repeating the analysis for gross (i.e. before fee) returns, the above results apply a fortiori. The implication here is that positive net-alpha performance depends crucially on competition amongst funds which could lead to lower mutual fund fees for investors in active funds. This is the focus of the UK Financial Conduct Authority's Asset Management Markets investigation (FCA 2017) which recommended greater transparency of fund performance measures, investigation into (internet) platform providers for funds and a submission to the Competition and Markets Authority (CMA) to investigate the role of advisers in recommending mutual funds to pension trustees.

The rest of the article is structured as follows. Section "Previous studies" provides a review of recent issues and empirical findings from the literature. In section "data", we discuss our data and sources while our empirical findings on performance persistence for UK funds are presented in Section "Empirical Results". Conclusions are presented in Section 5.

Previous studies

Factor models applied to mutual funds (MFs) are either based on "academic risk factors" which have been shown to explain the cross section of average stock returns or on "practitioner factors" which aim to represent low-cost investible passive indices. The problem in using academic factor models is the large number of potential alternative factors and the lack of a preferred model that price alternative stock portfolios (see for example, Harvey et al. 2015; Fama and French 2018).

The most frequently used academic factor model in the MF performance literature is the Fama–French–Carhart four-factor model (FFC4, Carhart 1997), with more recent work using bootstrap techniques to take account of non-normality in fund returns (Kosowski et al. 2006, Cuthbertson et al. 2008, Fama–French 2010, Blake et al. 2017, Huang et al. 2019,).

The evidence for both the UK and USA based on the FFC4 model is that the *average* MF underperforms in terms of (after-fee) net alpha. The evidence on *individual* fund performance is that positive net-alpha performance is relatively small but negative-alpha funds are much more prevalent, especially after adjusting for the false discovery rate or applying bootstrap p values[1]. Other studies for the USA and UK reinforce the above conclusions (e.g. For the USA, see inter alia, Kosowski et al. 2006, Fama and French 2010, Cai et al. 2018, Huang et al. 2019 and for the UK see Blake and Timmermann 1998, Quigley and Sinquefield 2000, Fletcher and Forbes 2002, Tonks 2005, Keswani and Stolin 2008, Cuthbertson et al. 2008, 2010, Mateus et al. 2019a, b)[2]. However, in a wide-ranging survey of US fund performance, Cremers (2017) and Cremers et al. (2019) describe the finding of few positive US alpha funds as the "conventional wisdom"—but they argue that recent studies suggest that the conventional wisdom is "too negative".

Testing for performance persistence

In testing for performance persistence, funds are typically sorted into fractiles (e.g. deciles) based on an attribute under examination (e.g. past performance) and periodically rebalanced over a specific holding period (e.g. monthly). Post-sort returns are then used to assess future performance. Alphas can be measured using *gross fund returns*[3] while investors earn *net returns* (i.e. gross returns after deduction of fund management fees)[4]. This provides a direct test of whether a particular ex-ante strategy could have been successful (on past data), for investors switching between funds[5].

[1] Andrikogiannopoulou and Papakonstantinou (2019) using simulation, show that for US mutual fund returns, which have a low signal-to-noise ratio, relatively limited observations per fund and possible cross-sectional correlation across funds, estimates of the false discovery rate may be heavily biased and produce estimates of zero-alphas (non-zero-alphas) that are upward (downward) biased. Giglio et al (2018) propose an alternative FDR approach.

[2] Another major strand in the literature is to use some form of shrinkage to adjust individual fund alphas. This may involve using Bayesian priors (Jones and Shanken 2005) or fund characteristics (Pastor and Stambaugh 2002) or information on the cross-section of performance (Chen et al 2017, Harvey and Liu 2018, 2019).

[3] After transactions costs of buying and selling securities but before deduction of management fees.

[4] Net returns exclude any load fees and any income or capital gains taxes applicable to the individual investor.

[5] We do not include switching costs in our analysis, which potentially include front and back-end load fees and time and effort of investors

Measuring persistence in MF performance has used a wide variety of sorting rules. For example, fund sorts based on measures of past performance such as fund returns, benchmark-adjusted returns, alphas and t-alphas are often used[6]. But other sorting rules include fund size (Cremers and Pareek 2016), active share (Cremers and Petajisto 2009; Frazzini et al. 2016; Petajisto 2016), industrial concentration (Kacperczyk et al. 2005), turnover (Pastor et al. 2017), "unobserved actions" (Kacperczyk et al. 2008), R-squared (Amihud and Goyenko 2013), "word of mouth" (Hong et al. 2005), "commonality in holdings" (Cohen et al. 2005), public information (Kacperczyk and Seru 2007), patient capital, (Cremers and Pareek 2016), gross profitability (Kenchington et al. 2019), efficiency of trading desks (Cici et al. 2018), herding behaviour (Jiang and Verardo 2018) and past cash inflows (i.e. the smart money effect, Zheng 1999, Sapp and Tiwari 2004, Keswani and Stolin 2008, Akbas et al. 2015)[7]. These are usually single sorts but sometimes double sorts on two attributes are used[8]. As the number of possible rules for predicting fund returns is very large, issues of data mining and data snooping come to the fore.

Using the FFC4 model on US data, Kosowski et al. (2006), for example, demonstrate that the top decile exhibits persistence with annual rebalancing to give a net alpha of 1% p.a. (bootstrap p value = 0.05). At the bottom of the performance distribution, deciles 6–10 have significantly negative abnormal performance (of about −1% p.a. for deciles 6–9 and −3.5% for decile 10).

More recent US studies (Jordan and Riley 2015, 2016) have added the Fama and French (2015, 2016, 2017) pricing factors for profitability (RMW, "robust minus weak") and investment (CMA, "conservative minus aggressive") to the FFC4 four-factor model, giving a six-factor model (FFC6). Jordan and Riley (2015) find that sorting MFs on their previous year's return volatility (and rebalancing each year) gives a positive (net of fees) alpha when using the FFC4 model, but the alpha "disappears" when using the FFC6 model. Hence, the FFC6 model negates the so-called volatility anomaly found in the US asset pricing literature (Frazzini and Pederson 2013, Novy-Marx 2014; Fama and French 2015, 2016).

Benchmark-adjusted alphas

Mateus et al. (2019a, b) suggest two methods to adjust academic factor models for the presence of nonzero alphas, when a passive benchmark (e.g. FTSE100, S&P500) is the dependent variable. The first approach looks for "new" factors that explain the cross section of stock returns but they argue that there is no consensus set of factors that emerges and any "new" factors added to the FFC4 model do not make a substantive difference to our views on MF alpha performance. A second method is to use as the dependent variable the fund's return minus the return on the fund's self-declared benchmark return. Then, the fund's true alpha is given by $\alpha = \alpha_{FFC4} - \alpha_b$. Clearly, if the MF benchmark-alpha is positive (negative), then the true alpha is smaller (larger) than the MFs standard FFC4 alpha. Mateus et al. (2016) find that for UK funds (1992–2013) adjustment for a benchmark return (i.e. FTSE100) tends to increase *historical* (Morningstar) style-alphas, compared with the FFC4 model. Applying a similar procedure, Mateus et al. (2019a, b) add a "peer group style-benchmark" to the FFC4 model and find that UK (Morningstar) peer group alphas increase the probability that these funds will be in the top-performing quartile, one year later. Of course, the choice of a specific benchmark may not adequately represent a mutual fund's risk exposure (Sensoy 2009).

Index models

As we have seen above, a potential problem in using the FFC4 academic factors is that mainstream *passive portfolios* (e.g. S&P500, Russell 2000, FTSE100) give nonzero alphas (see Mateus et al. 2016 for the UK and Cremers et al. 2013 for the US). Investment practitioners interpret alpha primarily in terms of "outperformance" relative to passive "index style-factors" that can either easily be replicated by individual investors themselves or by purchasing index funds directly (Sharpe 1992). Unlike the Fama–French academic factors, these index factors (IDX4, IDX7) generally have zero alphas with respect to a wide range of passive indices and style-sorted stock portfolios (e.g. stocks sorted on size or BMV) and also have higher R-squared and lower tracking errors than the FFC4 model (Cremers et al. 2013).

Footnote 5 (continued)

and advisers. For switches within fund families or for large investors such as pension funds actual load fees paid will be less than advertised load fees. Not all funds have load fees.

6 Past returns, alphas or t-alphas are perhaps the most widely used sorting rules in early work (e.g. Carhart 1997, Blitz and Huij 2012) and their use continues to today. Blake and Morey (2000) sort on Morningstar 5-star ratings.

7 In some of these studies only results for the future alphas (or returns) of the long-short sorted portfolios (i.e. top fractile funds minus bottom fractile funds) are reported. As mutual funds cannot be shorted, it is not clear if such results are exploitable by investors.

8 For example, Bessler et al. (2018) sort on the size of fund inflows ("external governance") and change in fund manager ("internal governance") while Cremers and Petajisto (2009) sort on active share and fund size.

DATA

The UK data set contains 779 UK equity mutual funds taken from Morningstar. Funds are categorised as small-cap, mid-cap and large-cap and as value, income, growth and blend funds. Index tracking funds are excluded. There are 334 non-surviving funds in the sample. Of the total 779 UK funds, 159 are domiciled offshore. Returns are total monthly returns net of management fees. We include the oldest share class of each fund.

Our sample period is from January 1990 to February 2017. In January 1990 there is a low of 157 funds, while in March 2009 there is a high of 600 funds. The average month has 391 funds in the sample. The cross-sectional average fund return ranges from a low of -15.28% in October 2008 to a high of 12.35% in April 2009[9]. The time series average of the monthly cross-sectional average fund return is 0.42%. At the end of the sample period, the average fund size was £510m.

In our factor model estimation, we source the Fama–French–Carhart market, size, value and momentum factors for the UK from the Xfi Centre for Finance and Investment data library at the University of Exeter Business School[10]. The factor construction methodology is outlined in Gregory et al. (2013). In the UK four-factor index model (UK-IDX4), the factor returns are the FTSE All Share (FA), FTSE Small (FS) minus FTSE 100 (F100), FTSE Value (FV) minus FTSE Growth (FG) indices and the momentum factor. The indices are sourced from Datastream. Due to reduced data availability, these data cover the period January 1994–February 2017. Statistical significance (based on bootstrap p values) is reported at the 5% significance level.

We implement a nonparametric bootstrap procedure to take account of non-normality in the residuals of the factor model used, which is particularly acute for funds in the tails of the performance distribution that we are particularly interested in. Bootstrapped p values of the t-statistic are generated as follows. For each fund, $i = 1,2..N$, over the period T_i, we bootstrap fund excess returns under the null hypothesis of zero true alpha. Amongst this set of N simulated returns, there is no persistence by construction as all funds have the same zero true alpha. Using these bootstrap simulated returns, we sort funds into "past winners and losers" and track the forward looking (holding period) "persistence" returns as we rebalance each month. We then estimate the t-statistic on the "persistence-alpha" for the top decile (or smaller size portfolios), under the null of a zero alpha. We repeat the above over a thousand bootstrap iterations to generate a distribution of this "persistence" t-statistic under the null of a zero "persistence-alpha".

Bootstrap p values are then calculated by comparing the empirical (actual) t-statistic on alpha from the historical data, with the null distribution for t-alpha from the simulated returns data under the null. For example, if the 95% cut-off amongst the 1000 simulated t-alphas (under the null of zero alpha) is 2.5 and the empirical t-alpha from the historical data is 2.2, then we do not reject the null of a zero alpha when using the bootstrap procedure (whereas we would reject the null under the assumption of normality) A key element of the bootstrap procedure is that we do *not assume* the estimated alpha for each fund is normally distributed. Each fund's empirical alpha can follow any distribution (depending on the fund's residuals) and this distribution can be different for each fund.

Empirical results

Based on the above analysis, for our UK funds we use the FFC4 model as our "academic factor" model and the IDX4 model as our "practitioner index" model. In this section, we present the empirical results on persistence in performance for both types of model. For the most part, we focus on net returns and net alphas (unless otherwise stated).

Decile-sorted portfolios: (f, h)= (60,1)

We begin by considering decile portfolios for (f, h)= (60,1) and net (after fees) returns in Table 1. First, the factor loadings on both the market risk factor and size risk factor are statistically significant for all deciles as well as for a portfolio of decile 1 minus decile 10, denoted "Decile 1-10". The value factor (HML) is also significant for decile 1, while the momentum risk factor (UMD) is significant for the top three deciles at the 5% significance level. Value and momentum are also significant for decile 1-10. In order to avoid presenting an overly large volume of tables and results, we concentrate throughout this section on the FFC4 model and the IDX4 index model. The FFC4 model statistically dominates the FF3 factor model (because of the statistical significance of the momentum factor, UMD) and results for IDX4 and IDX7 are qualitatively similar[11]. As with previous UK studies, for the FFC4 factor model there is some evidence of top-decile positive net-alpha persistence (α=2.12%pa, t=2.86) and bottom-decile negative net-alpha persistence (α= −1.37%pa, t = −2.93, Table 1).

[9] In April 2009 the UK stock market (e.g. FTSE All Share index) rose by 10% following a London G20 summit where an agreement was reached to tackle the global financial crisis with measures worth $1.1 trillion (£681bn).

[10] Available at http://business-school.exeter.ac.uk/research/centres/xfi/famafrench.

[11] The full set of results are available on request.

Table 1 UK Performance Persistence: Decile Portfolios—Fama–French–Carhart Four-Factor (FFC4) Model, (f,h = 60,1) This table presents the performance persistence results of decile-sorted mutual funds. Each month funds are sorted into equally weighted decile portfolios based on the t-statistic of alpha from a Fama–French–Carhart four-factor (FFC4) model estimated over the previous 60 months formation period. Each decile portfolio is held for a one-month holding period and the process is repeated on a one-month rolling basis. A time series of holding period returns is generated for each decile and the FFC4 model is estimated in each case over the holding period returns. The table shows the alpha (annualised) and factor loadings (betas) for each of these decile regressions. Also shown are the alpha and betas of (i) an equally weighted portfolio of all funds, denoted "All Funds" and (ii) a portfolio of the top decile minus the bottom decile of funds, denoted "Decile 1 – 10". t-statistics are shown in parentheses. Newey–West adjusted t-statistics are calculated throughout. Rm, SMB, HML and UMD refer to the market, size, value and momentum risk factors, respectively. Rm is measured in excess of the risk free rate. Also shown are the R^2 values as well as the Jarque–Bera (JB) test statistic of the null hypothesis that the regression residuals follow a normal distribution (p values in brackets). Results relate to the sample period January 1990–February 2017. Funds with a minimum of 60 observations are used leaving 660 funds

Portfolio	All Funds	Decile 1	Decile 2	Decile 3	Decile 4	Decile 5	Decile 6	Decile 7	Decile 8	Decile 9	Decile 10	Decile 1-10
Alpha	0.009	2.124	0.696	0.756	0.600	−0.078	0.504	−0.624	−0.624	−0.732	−1.368	3.492
	(0.020)	(2.860)	(1.046)	(1.302)	(0.864)	(−0.134)	(0.678)	(−1.038)	(−1.031)	(−1.210)	(−2.928)	(6.509)
Rm	0.932	0.901	0.917	0.919	0.931	0.945	0.935	0.937	0.931	0.948	0.957	−0.056
	(86.712)	(57.088)	(50.854)	(64.376)	(60.521)	(57.420)	(54.343)	(55.381)	(51.649)	(55.794)	(68.900)	(−4.957)
SMB	0.282	0.399	0.352	0.335	0.272	0.272	0.245	0.229	0.206	0.190	0.146	0.252
	(16.446)	(15.342)	(16.229)	(18.910)	(12.382)	(11.688)	(11.000)	(7.546)	(7.456)	(7.898)	(7.963)	(19.410)
HML	0.022	0.067	0.023	−0.017	0.017	0.036	0.034	0.016	0.018	0.004	0.005	0.062
	(1.473)	(2.738)	(0.907)	(−0.929)	(0.839)	(1.816)	(1.850)	(0.676)	(0.806)	(0.197)	(0.338)	(4.211)
UMD	0.021	0.029	0.035	0.027	0.015	0.020	0.015	0.018	0.015	0.011	0.006	0.023
	(1.573)	(1.655)	(1.897)	(1.703)	(0.916)	(1.426)	(0.991)	(1.079)	(0.958)	(0.647)	(0.514)	(2.258)
R^2	0.971	0.944	0.950	0.961	0.954	0.958	0.946	0.950	0.951	0.955	0.966	0.599
JB	24.7	51.3	70.5	33.3	28.2	20.0	937.1	57.6	34.7	34.9	11.0	26.1
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.003)	(0.000)

In the case of 69% of the 660 funds, we reject the null hypothesis of normally distributed residuals in the FFC4 model

When using the IDX4 index model, the above persistence results are broadly similar (Table 2), i.e. α=2.22%pa, t = 2.04 for the top decile and a statistically significant negative alpha for the bottom decile α=−2.06% p.a. t = −1.99.

Small portfolios: (*f*, *h*)= (60,1)

Next, we zoom in on the top decile of fund performance and examine alternative small-size portfolios of UK past winner funds within the more extreme tail of the cross section of fund performance, i.e. small-size portfolios s_i (i = 2, 3, 5, 7,10,20, 35, 50). For now, we present results for (f, h) = (60,1) in Table 3 and Table 4. We find strong evidence of positive (net) alpha persistence for all s_i for the FFC4 model (Table 3). For $s_i \leq 10$, the alphas are larger than the decile-1 alpha in Table 1 (i.e. for the same model). In Table 3, in the case of alpha we also report the bootstrap p value of the t-statistic of alpha in square brackets. Significance findings are consistent between the conventional t-statistic and the nonparametric bootstrap p value. Similarly, for the IDX4 index model results (Table 4), for small-size winner portfolios (with $s_i \leq 20$), statistically significant (net) alphas are in the range 2.8%—4.8% pa and are all larger than the decile-1 alpha in Table 2 (for the same model). There is a monotonic decline in alphas as the size of the winner portfolio increases from 2 to 50 funds (Table 4). Amongst the significant alphas for $s_i \leq 20$, significance findings are consistent between the conventional t-statistic and the nonparametric bootstrap p value. However, for $s_i \geq 20$ significance tests are sensitive to the choice between conventional t-statistic versus bootstrap p value where the t-statistic indicates significance at the 5% significance level but the bootstrap p value does not. Also, in contrast to US results (Cuthbertson et al. 2022), these small UK winner portfolios hold momentum stocks (see the UMD coefficients in Table 3 and Table 4).

Formation/holding periods: top-decile and top-small portfolios

When examining *decile portfolios* sorted on past t-alphas, out of the 8 possible combinations of (f, h), we find positive persistence for six *top-decile* UK funds for the FFC4 factor model but only for two top-decile funds for the IDX4 index model (Table 5, Panel A).

After sorting UK funds into small portfolios, there are 8 size portfolios x 8 combinations of (f, h). For these small-size winner portfolios, the FFC4 factor model indicates quite consistent findings across (f, h) where 53 (out of 64)

Table 2 UK Performance Persistence: Decile Portfolios—Four-Factor Index (IDX4) Model, (f,h = 60,1) This table presents the performance persistence results of decile-sorted mutual funds. Each month funds are sorted into equally weighted decile portfolios based on the t-statistic of alpha from the four-factor index (IDX4) model estimated over the previous 60 months formation period. Each decile portfolio is held for a one-month holding period and the process is repeated on a one-month rolling basis. A time series of holding period returns is generated for each decile and the IDX4 model is estimated in each case over the holding period returns. The table shows the alpha (annualised) and factor loadings (betas) for each of these decile regressions. Also shown are the alpha and betas of (i) an equally weighted portfolio of all funds, denoted "All Funds" and (ii) a portfolio of the top decile minus the bottom decile of funds, denoted "Decile 1 – 10". t-statistics are shown in parentheses. Newey–West adjusted t-statistics are calculated throughout. FA, FS, F100, FV and FG refer to the returns on the FTSE All Share, FTSE Small, FTSE 100, FTSE Value and FTSE Growth indices, respectively. FA is measured in excess of the risk free rate. UMD refers to the momentum risk factor. Also shown are the R^2 values as well as the Jarque–Bera (JB) test statistic of the null hypothesis that the regression residuals follow a normal distribution (p values in brackets). Results relate to the sample period January 1994–February 2017. Funds with a minimum of 60 observations are used leaving 660 funds

Portfolio	All Funds	Decile 1	Decile 2	Decile 3	Decile 4	Decile 5	Decile 6	Decile 7	Decile 8	Decile 9	Decile 10	Decile 1-10
Alpha	0.360	2.220	1.032	0.756	0.003	−0.768	−0.396	−0.708	−1.272	−1.896	−2.064	4.284
	(0.456)	(2.037)	(1.018)	(0.735)	(0.004)	(−0.771)	(−0.381)	(−0.694)	(−1.258)	(−1.905)	(−1.986)	(5.589)
FA	0.851	0.836	0.854	0.841	0.844	0.849	0.842	0.845	0.856	0.859	0.851	−0.015
	(31.651)	(26.577)	(27.763)	(25.814)	(27.814)	(25.427)	(23.628)	(24.400)	(25.743)	(24.862)	(24.568)	(−1.030)
FS-F100	0.375	0.556	0.470	0.418	0.370	0.368	0.327	0.309	0.318	0.343	0.329	0.226
	(10.726)	(12.529)	(11.688)	(10.183)	(8.877)	(8.480)	(6.248)	(6.212)	(6.864)	(7.916)	(6.845)	(11.980)
FV-FG	0.035	0.064	0.074	0.038	0.051	0.022	0.026	0.010	0.034	0.017	0.023	0.041
	(0.730)	(1.441)	(1.612)	(0.751)	(0.976)	(0.443)	(0.459)	(0.175)	(0.573)	(0.303)	(0.403)	(1.761)
UMD	0.048	0.074	0.073	0.048	0.036	0.031	0.026	0.019	0.027	0.044	0.022	0.052
	(1.807)	(2.225)	(2.267)	(1.512)	(1.293)	(1.105)	(0.676)	(0.610)	(0.982)	(1.747)	(0.654)	(4.182)
R^2	0.872	0.875	0.878	0.868	0.861	0.862	0.841	0.841	0.848	0.847	0.846	0.426
JB	115.6	37.3	69.0	94.2	90.3	47.4	76.2	56.9	42.2	86.9	54.9	0.606
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.738)

In the case of 80% of the 660 funds, we reject the null hypothesis of normally distributed residuals in the UK-IDX4 index model

combinations exhibit positive statistically significant alphas (Table 6, Panel A). Statistically significant inferences are also highly consistent whether we apply the conventional t-statistic or the nonparametric bootstrap p value where in all bar just 2 cases ($s_5, f, h = 36, 12$ and $s_{50}, f, h = 60, 6$) the findings with respect to statistical significance at the 5% significance level are qualitatively the same by both criteria. The IDX4 model (Table 7, Panel A) indicates widespread alpha persistence at the 5% significance level amongst most of the various small-size portfolios for $f = 60$, over all holding periods. However, there is considerably less evidence of persistence at the 5% significance level for shorter formation periods ($f = 36$) based on the bootstrap p value. Furthermore, over this shorter formation period, statistical significance is highly sensitive to the choice of standard t-statistic versus bootstrap p value, where the former supports significant persistence but the latter does not.

The statistically significant net alphas for small portfolios of funds in Table 6 (Panel B) and Table 7 (Panel B) are also economically significant ranging from 1.18% p.a. to 4.8% p.a. These net alphas tend to be larger for (i) relatively smaller portfolios (move down each column) and for (ii) shorter holding periods (move across each row). Also, alphas are generally larger when sorting using a longer formation period (i.e. f=60 rather than f =36).

Overall, the UK results show larger statistically significant net alphas for relatively small-size portfolios over long formation periods and short holding periods. Both the academic factor model and practitioner index model exhibit performance persistence although this finding is less robust (over all formation periods) in the case of the latter model, where statistically significant inferences are also sensitive to non-normality in fund returns. The UK results demonstrate robustness in alpha-performance persistence. when forming small past winner portfolios of funds. This is because even small portfolios of mutual funds are well diversified but comprise fewer funds with relatively low t-statistics compared to the larger decile portfolios.

Recursive estimates: FFC4 model[12]

Figure 1 shows recursive estimates of (net) alpha (brown line) and t-alpha (blue line) for the FFC4 model for a

[12] In the interests of brevity we report these robustness variants only for the FFC4 model but results using the IDX4 model are qualitatively similar and are available on request.

Table 3 UK Performance Persistence: Small Portfolios—Fama–French–Carhart Four-Factor (FFC4) Model, (f,h = 60,1) This table presents the performance persistence results of portfolios of mutual funds sorted by various size portfolios. Each month funds are sorted into equally weighted portfolios of size 2,3,5,7,10,20,35 and 50 based on the t-statistic of alpha from a Fama–French–Carhart four-factor (FFC4) model estimated over the previous 60 months formation period. Each size portfolio is held for a one-month holding period and the process is repeated on a one-month rolling basis. A time series of holding period returns is generated for each size portfolios and the FFC4 model is estimated in each case over the holding period returns. The table shows the alpha (annualised) and factor loadings (betas) for each of these size portfolios regressions. t-statistics are shown in parentheses. Newey–West adjusted t-statistics are calculated throughout. In the case of alpha, we also report the bootstrap p value of the t-statistic of alpha in square brackets. Rm, SMB, HML and UMD refer to the market, size, value and momentum risk factors, respectively. Rm is measured in excess of the risk free rate. Also shown are the R^2 values as well as the Jarque–Bera (JB) test statistic of the null hypothesis that the regression residuals follow a normal distribution (p values in brackets). Results relate to the sample period January 1990–February 2017. Funds with a minimum of 60 observations are used leaving 660 funds

Portfolio	Size 2	Size 3	Size 5	Size 7	Size 10	Size 20	Size 35	Size 50
Alpha (t-stat) [p value]	4.152	4.572	3.864	2.844	2.436	2.100	2.040	1.752
	(2.972)	(3.693)	(3.676)	(3.032)	(2.804)	(2.686)	(2.803)	(2.670)
	[0.003]	[0.000]	[0.000]	[0.001]	[0.005]	[0.006]	[0.001]	[0.004]
Rm	0.867	0.878	0.893	0.907	0.891	0.904	0.911	0.910
	(22.839)	(28.611)	(33.080)	(39.574)	(45.787)	(52.231)	(52.700)	(57.692)
SMB	0.457	0.476	0.468	0.439	0.423	0.404	0.403	0.394
	(11.103)	(12.863)	(16.457)	(15.390)	(16.056)	(15.521)	(15.521)	(18.097)
HML	0.221	0.173	0.127	0.111	0.097	0.076	0.049	0.042
	(5.491)	(5.278)	(4.529)	(4.442)	(4.525)	(3.135)	(1.794)	(1.880)
UMD	0.048	0.050	0.043	0.038	0.041	0.034	0.029	0.037
	(1.648)	(2.071)	(1.896)	(1.779)	(2.105)	(1.874)	(1.638)	(2.251)
R^2	0.828	0.867	0.902	0.914	0.926	0.940	0.944	0.951
JB	1.247	5.767	3.342	15.3	19.9	35.1	86.6	86.2
	(0.536)	(0.055)	(0.179)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

portfolio comprising the *top-5* UK funds, for various combinations of formation and holding periods (f,h). Over time, alphas are more statistically significant for longer (f=60) versus shorter formation periods (f=36) and for shorter versus longer holding periods, h. Alphas are also more constant over time using longer formation periods (i.e. f=60 versus f=30).

We now take (f, h) = (60,1) for the FFC4 model and examine its performance as we alter the number of funds in the portfolio. Figure 2 shows that smaller size portfolios ($s \leq 5$ funds) yield larger and more statistically significant persistence-alphas over time. These recursive estimates support the use of relatively small-size portfolios, formed using relatively long formation periods and held over relatively short holding periods.

Net versus gross alphas

In moving from UK investors' net alphas to gross alphas (which measure persistence before deduction of fees), there is little qualitative change in the results. In the case of both the FFC4 academic factor model and the IDX4 practitioner index model, the decile gross alphas (Table 8) are larger than the corresponding decile net alphas (Table 5) over all values of (f, h). In addition, in the case of both models, there are also more statistically significant decile gross alphas than decile net alphas in total, over the alternative values of (f, h).

Similarly, in the case of size portfolios s_i (i = 2, 3, 5, 7,10,20, 35, 50), for both the FFC4 model (Table 9) and IDX4 model (Table 10), the gross alphas are larger than the corresponding net alphas (Table 6 and Table 7) in almost all cases over the alternative values of (f, h). Indeed, when sorting into 8 small-size fund portfolios s_i (i = 2, 3, 5, 7,10,20, 35, 50) and examining the 8 alternative combinations of (f, h), we find that nearly all 64 combinations have statistically and economically significant gross alphas for both the FFC4 (Table 9) and IDX4 (Table 10) models. Again, this reinforces the point that using small-size

Table 4 UK Performance Persistence: Small Portfolios—Four-Factor Index (IDX4) Model, (f,h = 60,1) This table presents the performance persistence results of portfolios of mutual funds sorted by various size portfolios. Each month funds are sorted into equally weighted portfolios of size 2,3,5,7,10,20,35 and 50 based on the t-statistic of alpha from the four-factor index (IDX4) model estimated over the previous 60 months formation period. Each size portfolio is held for a one-month holding period and the process is repeated on a one-month rolling basis. A time series of holding period returns is generated for each size portfolios and the IDX4 model is estimated in each case over the holding period returns. The table shows the alpha (annualised) and factor loadings (betas) for each of these size portfolios regressions. t-statistics are shown in parentheses. Newey–West adjusted t-statistics are calculated throughout. In the case of alpha, we also report the bootstrap p value of the t-statistic of alpha in square brackets. FA, FS, F100, FV and FG refer to the returns on the FTSE All Share, FTSE Small, FTSE 100, FTSE Value and FTSE Growth indices, respectively. FA is measured in excess of the risk free rate. UMD refers to the momentum risk factor. Also shown are the R^2 values as well as the Jarque–Bera (JB) test statistic of the null hypothesis that the regression residuals follow a normal distribution (p values in brackets). Results relate to the sample period January 1994–February 2017. Funds with a minimum of 60 observations are used leaving 660 funds

Portfolio	Size 2	Size 3	Size 5	Size 7	Size 10	Size 20	Size 35	Size 50
Alpha (t-stat) [p value]	4.836	4.476	4.320	3.804	3.732	2.868	2.004	1.860
	(2.779)	(2.781)	(2.926)	(2.708)	(2.864)	(2.429)	(1.846)	(1.782)
	[0.009]	[0.006]	[0.003]	[0.010]	[0.002]	[0.014]	[0.054]	[0.074]
FA	0.785	0.795	0.813	0.819	0.824	0.834	0.848	0.850
	(16.087)	(19.276)	(21.413)	(21.999)	(22.539)	(24.568)	(27.268)	(27.603)
FS-F100	0.571	0.635	0.660	0.674	0.660	0.590	0.567	0.538
	(7.235)	(10.135)	(12.565)	(14.080)	(15.746)	(12.271)	(13.028)	(12.529)
FV-FG	0.123	0.104	0.092	0.066	0.070	0.067	0.073	0.063
	(1.850)	(1.763)	(1.829)	(1.430)	(1.599)	(1.451)	(1.597)	(1.328)
UMD	0.041	0.064	0.082	0.089	0.100	0.076	0.080	0.078
	(0.755)	(1.708)	(2.404)	(2.835)	(3.607)	(2.229)	(2.639)	(2.574)
R^2	0.742	0.797	0.843	0.854	0.863	0.869	0.880	0.881
JB	59.3	9.92	25.4	26.4	15.6	14.7	18.8	37.2
	(0.000)	(0.006)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Table 5 UK Performance Persistence: Top Decile—various formation/holding periods. This table presents the performance persistence results of mutual funds sorted by the top decile. Results relate to the Fama–French–Carhart four-factor (FFC4) model and the four-factor index (IDX4) model. At time *t*, funds are sorted into an equally weighted top-decile portfolio based on the t-statistic of alpha from the performance model estimated over the previous *f* months formation period. The top-decile portfolio is held for a *h* months holding period and the process is repeated on a *h* months rolling basis. A time series of holding period returns is generated for the top decile and the model is then estimated over the holding period returns. In Panel A, in each cell of the table we report the t-statistic of alpha and the bootstrap p value of the t-statistic as "t-stat|p value". Newey–West adjusted t-statistics are calculated throughout. Results are presented for alternative values of *f-h* as indicated. Panel B reports the corresponding annualised alphas. In the case of the FFC4 model, the sample period is January 1990–February 2017. In the case of the IDX4 model the sample period is January 1994–February 2017. When funds with a minimum of 36 (60) observations are used, there are 720 (660) funds in the analysis.

PANEL A: t-alpha | p value f-h

	36-1	36-3	36-6	36-12	60-1	60-3	60-6	60-12
FFC4	2.757|0.007	2.260|0.020	1.431|0.077	2.202|0.018	2.860|0.005	2.626|0.003	1.521|0.059	2.008|0.028
IDX4	1.649|0.132	1.575|0.153	1.055|0.327	1.215|0.298	2.037|0.035	2.069|0.030	0.914|0.276	1.648|0.089

PANEL B: Alpha (% p.a.) f-h

	36-1	36-3	36-6	36-12	60-1	60-3	60-6	60-12
FFC4	1.836	1.524	1.02	1.488	2.124	1.92	1.164	1.536
IDX4	1.416	1.452	1.104	1.248	2.220	2.328	1.188	1.776

portfolios of past winner funds rather than decile portfolios reveals more positive persistence (before deduction of fees). Finally, the analysis of fund gross returns indicates more widespread persistence compared to fund net returns over the values of (f, h). This suggests that in some cases performance persistence is achieved by, but is also absorbed by,

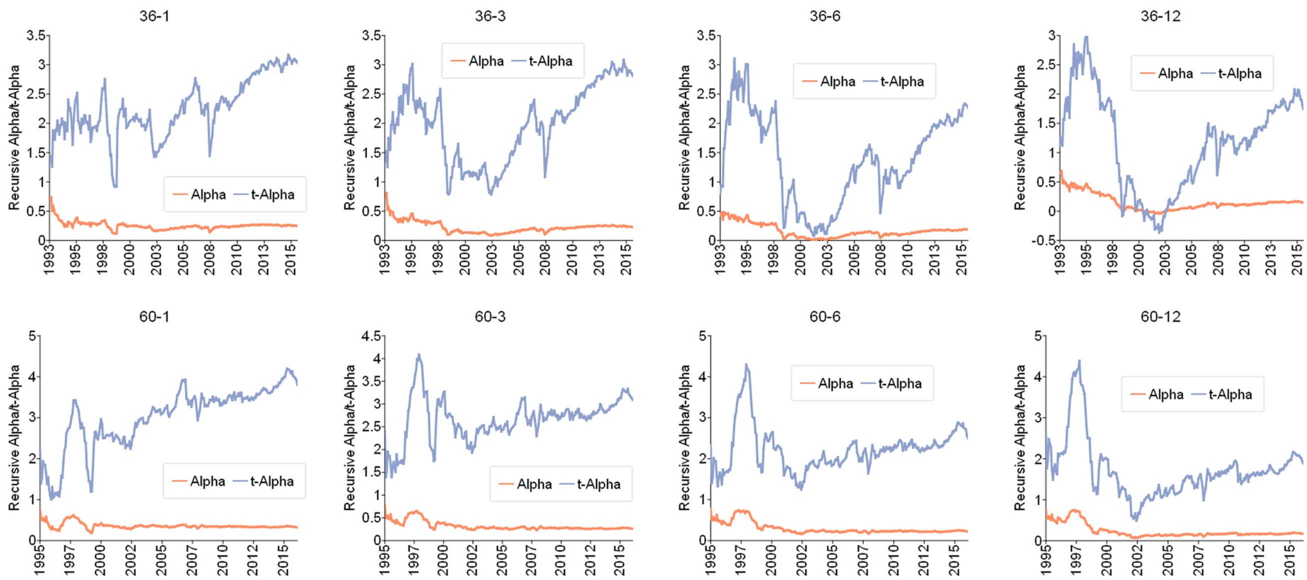

Fig. 1 Recursive Estimation of UK Top-5 Funds: Fama–French–Carhart Four-Factor (FFC4) Model, alternative f,h periods. The figure plots the recursive estimates of alpha and t-statistic of alpha of portfolios of the top 5 funds. Each month funds are sorted into equally weighted portfolios of size 5 funds based on the t-statistic of alpha from a Fama–French–Carhart four-factor (FFC4) model estimated over formation periods of 36 months and 60 months. Each portfolio is held for holding periods of 1,3,6 and 12 months and the process is repeated on a 1,3,6 and 12 month rolling basis, respectively. A time series of holding period returns is generated in each case. We then estimate the model recursively over the holding period returns. Each panel plots the recursive estimates of alpha and t-statistic of alpha for the formation and holding periods as indicated. Sample period 01/1990–02/2017

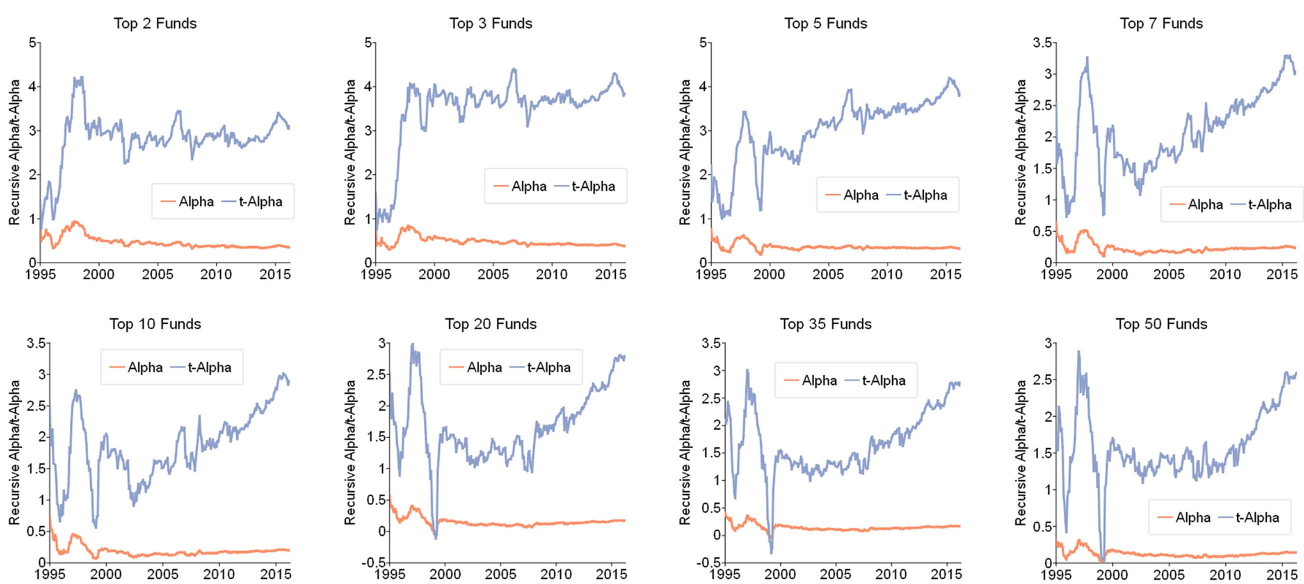

Fig. 2 Recursive Estimation of UK Size Portfolios : Fama–French–Carhart Four-Factor (FFC4) Model, (f,h = 60,1) The figure plots the recursive estimates of alpha and t-statistic of alpha of portfolios of mutual funds sorted by various size portfolios. Each month funds are sorted into equally weighted portfolios of size 2,3,5,7,10,20,35 and 50 based on the t-statistic of alpha from a Fama–French–Carhart four-factor (FFC4) model estimated over the previous 60 months formation period. Each size portfolio is held for a one-month holding period and the process is repeated on a one-month rolling basis. A time series of holding period returns is generated for each size portfolio. We then estimate the model recursively over these holding period returns. Each panel plots recursive estimates of alpha and t-alpha for the various size portfolios as indicated. Sample period 01/1990–02/2017

Table 8 UK Performance Persistence (Gross Returns): Top Decile—various performance models and formation/holding periods. This table presents the performance persistence results of mutual funds sorted by deciles based on fund gross returns. Results relate to alternative performance models as follows: Fama–French–Carhart four-factor (FFC4) model and the four-factor index (IDX4) model. At time *t*, funds are sorted into an equally weighted top-decile portfolio based on the t-statistic of alpha from the performance model estimated over the previous *f* months formation period. The top-decile portfolio is held for a *h* months holding period and the process is repeated on a *h* months rolling basis. A time series of holding period returns is generated for the top decile and the model is then estimated over the holding period returns. In Panel A, in each cell of the table we report the t-statistic of alpha and the bootstrap p value of the t-statistic as "t-stat|p value". Newey–West adjusted t-statistics are calculated throughout. Results are presented for alternative values of *f-h* as indicated. Panel B presents the corresponding annualised alphas. The sample period is January 1990–February 2017. When funds with a minimum of 36 (60) observations are used, there are 657 (607) funds in the analysis.

Panel A: t-alpha | p value t f-h

	36-1	36-3	36-6	36-12	60-1	60-3	60-6	60-12
FFC4	4.258|0.000	3.593|0.000	3.228|0.000	4.109|0.000	4.611|0.000	4.426|0.000	3.356|0.000	3.615|0.000
IDX4	3.386|0.002	3.101|0.008	2.368|0.028	2.402|0.035	3.255|0.002	3.266|0.002	2.108|0.030	2.795|0.008

Panel B: Alpha (% p.a.)

f-h	36-1	36-3	36-6	36-12	60-1	60-3	60-6	60-12
FFC4	2.909	2.465	2.152	2.663	3.276	3.136	2.506	2.668
IDX4	3.056	2.994	2.520	2.429	3.539	3.481	2.534	2.923

Table 10 UK Performance Persistence (Gross Returns): Small Portfolios—IDX4 index model—various formation/holding periods. This table presents the performance persistence results of portfolios of mutual funds sorted by various size portfolios based on fund gross returns. At time *t*, funds are sorted into equally weighted portfolios of size 2,3,5,7,10,20,35 and 50 based on the t-statistic of alpha from the four-factor index (IDX4) model estimated over the previous *f* months formation period. Each size portfolio is held for a *h* months holding period and the process is repeated on a *h* months rolling basis. A time series of holding period returns is generated for each size portfolio and the IDX4 model is then estimated over the holding period returns. In Panel A, in each cell of the table we report the t-statistic of alpha and the bootstrap p value of the t-statistic as "t-stat|p value". Newey–West adjusted t-statistics are calculated throughout. Results are presented for alternative values of *f-h* as indicated. Panel B reports the corresponding annualised alphas. Results relate to the sample period January 1994–February 2017. When funds with a minimum of 36 (60) observations are used, there are 720 (660) funds in the analysis.

Panel A: t-alpha | p value f-h

Size	36-1	36-3	36-6	36-12	60-1	60-3	60-6	60-12
2	2.141|0.027	2.920|0.006	2.431|0.018	2.074|0.040	2.627|0.007	2.117|0.028	1.790|0.043	2.320|0.021
3	2.402|0.015	2.753|0.004	2.488|0.016	1.971|0.057	2.944|0.002	3.396|0.000	3.068|0.001	2.795|0.004
5	2.680|0.006	2.679|0.009	2.000|0.052	2.407|0.021	3.430|0.002	3.457|0.000	3.149|0.002	3.303|0.001
7	3.125|0.003	2.822|0.007	2.268|0.027	2.180|0.036	3.365|0.000	3.703|0.000	3.137|0.003	3.176|0.002
10	3.514|0.002	3.301|0.005	3.012|0.006	2.297|0.045	3.322|0.001	3.329|0.002	2.795|0.008	2.791|0.009
20	3.327|0.003	3.124|0.005	3.015|0.006	2.868|0.011	3.280|0.000	3.354|0.001	2.701|0.010	3.097|0.001
35	3.341|0.004	3.212|0.003	2.560|0.020	2.570|0.030	2.920|0.004	3.069|0.003	2.028|0.039	2.648|0.015
50	3.085|0.010	2.990|0.011	2.270|0.054	2.624|0.028	2.709|0.008	2.511|0.014	1.691|0.083	2.207|0.031

Panel B: Alpha (% p.a.) f-h

Size	36-1	36-3	36-6	36-12	60-1	60-3	60-6	60-12
2	2.867	4.072	3.856	2.989	4.255	3.533	3.163	4.097
3	2.935	3.542	3.544	2.567	4.586	5.341	5.064	4.788
5	3.050	3.168	2.594	2.945	4.938	5.105	4.908	4.920
7	3.380	3.138	2.840	2.638	4.524	5.347	4.559	4.433
10	3.653	3.553	3.440	2.620	3.970	4.158	3.707	3.541
20	3.320	3.265	3.388	2.988	3.853	3.826	3.314	3.392
35	2.933	3.050	2.642	2.548	3.251	3.330	2.383	2.860
50	2.678	2.746	2.240	2.372	2.818	2.711	2.021	2.334

Table 6 UK Performance Persistence: Small Portfolios—FFC4 factor model–various formation/holding periods. This table presents the performance persistence results of portfolios of mutual funds sorted by various size portfolios. At time *t*, funds are sorted into equally weighted portfolios of size 2,3,5,7,10,20,35 and 50 based on the t-statistic of alpha from the Fama–French–Carhart four-factor (FFC4) model estimated over the previous *f* months formation period. Each size portfolio is held for a *h* months holding period and the process is repeated on a *h* months rolling basis. A time series of holding period returns is generated for each size portfolio and the FFC4 model is then estimated over the holding period returns. In Panel A, in each cell of the table we report the t-statistic of alpha and the bootstrap p value of the t-statistic as "t-stat|p value". Newey–West adjusted t-statistics are calculated throughout. Results are presented for alternative values of *f-h* as indicated. Panel B reports the corresponding annualised alphas. Results relate to the sample period January 1990–February 2017. When funds with a minimum of 36 (60) observations are used, there are 720 (660) funds in the analysis.

PANEL A: t-alphas | p values f-h

Size	36-1	36-3	36-6	36-12	60-1	60-3	60-6	60-12
2	2.089|0.023	1.212|0.123	1.500|0.081	0.613|0.286	2.972|0.003	2.913|0.001	2.293|0.008	0.696|0.316
3	2.724|0.007	2.226|0.022	2.024|0.030	1.397|0.088	3.693|0.000	3.405|0.000	2.238|0.016	1.022|0.193
5	3.068|0.002	2.751|0.007	2.154|0.018	1.703|0.055	3.676|0.000	2.944|0.001	2.390|0.006	1.783|0.043
7	2.880|0.004	3.409|0.000	3.477|0.001	3.146|0.000	3.032|0.001	2.551|0.009	1.836|0.037	1.239|0.131
10	2.734|0.004	2.905|0.002	2.930|0.003	3.421|0.000	2.804|0.005	2.578|0.009	1.725|0.049	1.546|0.074
20	2.188|0.016	2.209|0.014	1.909|0.022	2.746|0.005	2.687|0.006	2.311|0.011	0.976|0.168	1.915|0.045
35	2.807|0.003	2.167|0.015	1.726|0.042	2.104|0.023	2.804|0.001	2.376|0.015	1.493|0.066	2.305|0.013
50	2.812|0.000	2.250|0.009	2.116|0.018	2.386|0.013	2.670|0.004	2.331|0.010	1.610|0.047	1.865|0.026

PANEL B: Alphas (% p.a.)

Size	36-1	36-3	36-6	36-12	60-1	60-3	60-6	60-12
2	2.436	1.548	1.872	0.768	4.164	4.152	3.288	1.14
3	2.808	2.472	2.388	1.584	4.572	4.344	2.844	1.344
5	3.036	2.724	2.208	1.716	3.864	3.108	2.604	1.992
7	2.568	2.964	3.192	2.844	2.844	2.436	2.016	1.404
10	2.4	2.532	2.616	2.856	2.436	2.364	1.68	1.536
20	1.644	1.728	1.524	2.016	2.112	1.872	0.84	1.5
35	1.992	1.572	1.26	1.428	2.04	1.74	1.14	1.704
50	1.74	1.404	1.356	1.44	1.752	1.548	1.188	1.38

fund management companies rather than being enjoyed by investors[13].

Conclusions

As with previous studies, we find a statistical difference in the net alphas of top and bottom performing *decile-sorted* UK funds. As MFs cannot be short-sold, a long-short strategy cannot be implemented to exploit this difference in alphas and often the rather weak inference is for investors to avoid poor past performing funds. Hence, unless we can find persistent positive net-alpha funds, investors cannot benefit from a persistence strategy. In this paper, we therefore focus on tests of persistence in superior performance.

In contrast to studies of US equity mutual funds, we find more widespread evidence of UK decile portfolio persistence in net alphas across various formation and holding periods. This may simply be a feature of the relative sizes of the mutual fund industries where in the US case decile portfolios contain significantly more funds. When there are a large number of funds, the top decile may contain many funds with relatively small t-alphas. For the UK, there are fewer funds than for the USA so the top decile in the UK comprises fewer funds. This becomes much less of a problem if portfolios are formed from a smaller number of (past) top-performing funds. We find that sorting UK funds (by t-alpha) into relatively small portfolios (of up to 50 funds) reveals substantial positive persistence with

[13] We perform the fund sorting and holding analysis on net returns and gross returns *separately*. Therefore, the composition of funds in the holding period portfolios may not be the same in both cases. This may explain the differences in net alpha performance persistence compared to gross alpha performance persistence

Table 7 UK Performance Persistence: Small Portfolios—Four-Factor Index (IDX4) Model—various formation/holding periods. This table presents the performance persistence results of portfolios of mutual funds sorted by various size portfolios. At time *t*, funds are sorted into equally weighted portfolios of size 2,3,5,7,10,20,35 and 50 based on the t-statistic of alpha from the four-factor index (IDX4) model estimated over the previous *f* months formation period. Each size portfolio is held for a *h* months holding period and the process is repeated on a *h* months rolling basis. A time series of holding period returns is generated for each size portfolio and the IDX4 model is then estimated over the holding period returns. In Panel A, in each cell of the table we report the t-statistic of alpha and the bootstrap p value of the t-statistic as "t-stat|p value". Newey–West adjusted t-statistics are calculated throughout. Results are presented for alternative values of *f-h* as indicated. Panel B reports the corresponding annualised alphas. Results relate to the sample period January 1994–February 2017. When funds with a minimum of 36 (60) observations are used, there are 720 (660) funds in the analysis.

PANEL A: t-alphas |p values f-h

Size	36-1	36-3	36-6	36-12	60-1	60-3	60-6	60-12
2	1.680|0.065	1.803|0.063	1.796|0.067	1.693|0.082	2.779|0.009	2.066|0.041	1.677|0.063	1.300|0.114
3	1.943|0.032	2.058|0.052	1.975|0.059	1.891|0.056	2.782|0.006	2.669|0.007	2.337|0.010	1.921|0.034
5	2.184|0.024	2.091|0.048	1.847|0.087	1.832|0.081	2.926|0.003	2.867|0.002	2.361|0.013	2.201|0.021
7	2.574|0.011	2.528|0.020	1.953|0.068	1.695|0.101	2.708|0.010	2.771|0.007	2.115|0.030	2.031|0.031
10	2.458|0.011	2.403|0.026	1.829|0.083	1.459|0.169	2.864|0.002	2.598|0.009	2.063|0.027	2.218|0.020
20	1.672|0.089	1.901|0.094	1.728|0.106	1.754|0.122	2.429|0.014	2.459|0.012	1.739|0.070	1.700|0.084
35	2.057|0.056	2.069|0.065	1.489|0.204	1.431|0.229	1.846|0.054	2.037|0.034	0.976|0.260	1.779|0.066
50	1.879|0.101	1.856|0.117	1.092|0.372	1.187|0.322	1.782|0.074	1.611|0.107	0.812|0.334	1.493|0.128

PANEL B: Alphas (% p.a.)

Size	36-1	36-3	36-6	36-12	60-1	60-3	60-6	60-12
2	2.34	2.664	2.736	2.64	4.836	3.78	3.036	2.472
3	2.508	2.82	3.012	2.928	4.488	4.416	4.056	3.504
5	2.64	2.628	2.4	2.304	4.32	4.356	3.384	3.252
7	2.88	2.844	2.472	2.088	3.816	4.032	2.796	2.592
10	2.652	2.64	2.172	1.74	3.732	3.456	2.688	2.688
20	1.656	1.944	1.944	1.884	2.88	2.976	2.256	2.028
35	1.788	1.932	1.488	1.416	2.004	2.148	1.152	1.908
50	1.656	1.692	1.08	1.044	1.86	1.752	0.984	1.596

large statistically significant positive net alphas for many combinations (f, h).We have examined the robustness of this persistence in MF performance for UK equity MFs using monthly data (January 1990-February 2017) for different factor models, different size portfolios, different formation and holding periods (f, h) and using a cross-sectional bootstrap for tests on alpha.

Overall, our UK results provide support for positive net-alpha persistence for small portfolios of past winner funds (sorted on their t-alphas), with larger post-sort positive net alphas for formation periods based on 60 rather than 36 observations and for relatively short holding periods of up to 6 months. Hence, forming persistence portfolios with a small number of funds provides a relatively simple way for investors to obtain positive net alphas rather than searching over many alternative sorting rules or using multiple sorting criteria.

Table 9 UK Performance Persistence (Gross Returns): Small Portfolios—FFC4 factor model—various formation/holding periods. This table presents the performance persistence results of portfolios of mutual funds sorted by various size portfolios based on fund gross returns. At time *t*, funds are sorted into equally weighted portfolios of size 2,3,5,7,10,20,35 and 50 based on the t-statistic of alpha from the Fama–French–Carhart four-factor (FFC4) model estimated over the previous *f* months formation period. Each size portfolio is held for a *h* months holding period and the process is repeated on a *h* months rolling basis. A time series of holding period returns is generated for each size portfolio and the FFC4 model is then estimated over the holding period returns. In Panel A, in each cell of the table we report the t-statistic of alpha and the bootstrap p value of the t-statistic as "t-stat|p value". Newey–West adjusted t-statistics are calculated throughout. Results are presented for alternative values of *f-h* as indicated. Panel B presents the corresponding annualised alphas. Results relate to the sample period January 1990–February 2017. When funds with a minimum of 36 (60) observations are used, there are 720 (660) funds in the analysis.

Panel A: t-alpha | p value f-h

Size	36-1	36-3	36-6	36-12	60-1	60-3	60-6	60-12
2	2.608|0.011	2.565|0.005	2.678|0.007	2.010|0.032	2.673|0.004	2.444|0.009	2.378|0.016	1.034|0.171
3	3.450|0.000	3.667|0.000	3.829|0.000	2.641|0.013	3.484|0.000	2.846|0.005	2.665|0.007	2.247|0.020
5	3.668|0.000	3.888|0.000	4.139|0.000	3.788|0.001	4.151|0.000	4.162|0.000	3.304|0.002	3.396|0.001
7	3.126|0.002	3.327|0.000	4.167|0.000	4.292|0.000	3.434|0.002	3.635|0.000	2.779|0.003	3.559|0.000
10	3.281|0.002	3.651|0.000	3.912|0.000	3.644|0.000	4.109|0.000	3.964|0.000	2.797|0.003	3.44|0.000
20	3.566|0.000	3.163|0.000	3.456|0.001	3.740|0.000	4.657|0.000	4.226|0.000	2.812|0.002	3.071|0.000
35	4.247|0.000	3.692|0.000	3.229|0.000	3.906|0.000	4.613|0.000	4.537|0.000	3.394|0.000	3.958|0.000
50	4.327|0.000	4.018|0.000	3.706|0.000	4.162|0.000	4.571|0.000	4.411|0.000	3.524|0.000	3.712|0.000

Panel B: Alpha (% p.a.) f-h

Size	36-1	36-3	36-6	36-12	60-1	60-3	60-6	60-12
2	2.965	3.220	3.299	2.634	3.472	3.235	3.199	1.627
3	4.002	4.230	4.417	3.056	4.260	3.612	3.419	3.038
5	3.980	3.998	4.033	3.712	4.192	4.400	3.505	3.704
7	3.209	3.193	3.616	3.601	3.092	3.343	2.750	3.446
10	3.001	3.229	3.247	3.161	3.545	3.439	2.500	2.858
20	2.758	2.454	2.598	2.749	3.490	3.305	2.404	2.401
35	2.933	2.518	2.204	2.533	3.029	2.964	2.436	2.804
50	2.536	2.327	2.238	2.447	2.900	2.724	2.354	2.514

Acknowledgement We gratefully acknowledge financial support from the Irish Research Council (Project ID: IRC RFPS/2016/19).

Funding Open Access funding provided by the IReL Consortium.

Declarations

Conflict of interest The authors declare that there is no conflict of interest

References

Akbas, F., W.J. Armstrong, S. Sorescu, and A. Subrahmanyam. 2015. Smart money, dumb money and capital market anomalies. *Journal of Financial Economics* 118: 355–382. https://doi.org/10.1016/j.jfineco.2015.07.003.

Amihud, Y., and R. Goyenko. 2013. Mutual fund's R^2 as predictor of performance. *Review of Financial Studies* 26: 667–694. https://doi.org/10.1093/rfs/hhs182.

Andrikogiannopoulou, A., and F. Papakonstantinou. 2019. Reassessing false discoveries in mutual fund performance: skill, luck or lack of power?. *The Journal of Finance* 74 (5): 2667–2687. https://doi.org/10.1111/jofi.12784.

Bessler, W., D. Blake, P. Luckoff, and I. Tonks. 2018. Fund flows, manager changes and performance persistence. *Review of Finance* 22 (5): 1911–1947. https://doi.org/10.1093/rof/rfx017.

Blake, C., and M. Morey. 2000. Morningstar ratings and mutual fund performance. *Journal of Financial and Quantitative Analysis* 35 (3): 451–483. https://doi.org/10.2307/2676213.

Blake, D., and A. Timmermann. 1998. Mutual fund performance: evidence from the UK. *European Finance Review* 2: 57–77. https://doi.org/10.1023/A:1009729630606.

Blake, D., T. Caulfield, C. Ioannides, and I. Tonks. 2017. New evidence on mutual fund performance: A comparison of alternative bootstrap methods. *Journal of Financial and Quantitative*

Analysis 52 (3): 1279–1299. https://doi.org/10.1017/S0022109017000229.
Blitz, D. & Huij, J. (2012). Another look at the performance of actively managed equity mutual funds. SSRN 2004972.
Cai, B., T. Cheng, and C. Yan. 2018. Time-varying skills (versus luck) in US active mutual funds and hedge funds. *Journal of Empirical Finance* 49: 81–106. https://doi.org/10.1016/j.jempfin.2018.09.001.
Carhart, M. 1997. On persistence in mutual fund performance. *The Journal of Finance* 52 (1): 57–82. https://doi.org/10.1111/j.1540-6261.1997.tb03808.x.
Chen, Y., M. Cliff, and H. Zhoa. 2017. Hedge funds: The good, the bad and the lucky. *Journal of Financial and Quantitative Analysis* 52: 1081–1109. https://doi.org/10.1017/S0022109017000217.
Cici, G., L.K. Dahm, and A. Kempf. 2018. Trading efficiency of fund families: Impact on fund performance and investment behavior. *Journal of Banking and Finance* 88: 1–34. https://doi.org/10.1016/j.jbankfin.2017.11.004.
Cohen, R. B., Coval, J.D. & Pastor, L. (2005). Judging fund managers by the company they keep, *The Journal of Finance,* LX(3), 1057-96.
Cremers, M. 2017. Active share and the three pillars of active management: Skill, convictions and opportunity. *Financial Analysis Journal* 73: 61–79. https://doi.org/10.2469/faj.v73.n2.4.
Cremers, M., and A. Pareek. 2016. Patient capital outperformance: The investment skill of high active share managers who trade infrequently. *Journal of Financial Economics* 122: 288–306. https://doi.org/10.1016/j.jfineco.2016.08.003.
Cremers, M., and A. Petajisto. 2009. How active is your fund manager? A new measure that predicts performance. *Review of Financial Studies* 22 (9): 3329–3365. https://doi.org/10.1093/rfs/hhp057.
Cremers, M., A. Petajisto, and E. Zitzewitz. 2013. Should benchmark indices have alpha?. *Revisiting Performance Evaluation, Critical Finance Review* 2: 1–48. https://doi.org/10.1561/104.00000007.
Cremers, M., J.A. Fulkerson, and T.B. Riley. 2019. Challenging the conventional wisdom on active management: A review of the past 20 years of academic literature on actively managed mutual funds. *Financial Analyst Journal* 75 (4): 8–35. https://doi.org/10.1080/0015198X.2019.1628555.
Cuthbertson, K., D. Nitzsche, and N. O'Sullivan. 2008. UK mutual fund performance: skill or luck?. *Journal of Empirical Finance* 15 (4): 613–634. https://doi.org/10.1016/j.jempfin.2007.09.005.
Cuthbertson, K., D. Nitzsche, and N. O'Sullivan. 2022. Mutual fund performance persistence: factor models and portfolio size. *International Review of Financial Analysis*. https://doi.org/10.1016/j.irfa.2022.102133.
Fama, E.F., and K.R. French. 2010. Luck versus skill in the cross-section of mutual fund returns. *The Journal of Finance* 65: 1915–1947.
Fama, E.F., and K.R. French. 2015. A five-factor asset pricing model. *Journal of Financial Economics* 116: 1–22. https://doi.org/10.1016/j.jfineco.2014.10.010.
Fama, E.F., and K.R. French. 2016. Dissecting anomalies with a five-factor model. *Review of Financial Studies* 29: 69–103. https://doi.org/10.1093/rfs/hhv043.
Fama, E.F., and K.R. French. 2017. International tests of a five-factor asset pricing model. *Journal of Financial Economics* 123: 441–463. https://doi.org/10.1016/j.jfineco.2016.11.004.
Fama, E.F., and K.R. French. 2018. Choosing factors. *Journal of Financial Economics* 128 (2): 234–252. https://doi.org/10.1016/j.jfineco.2018.02.012.
Fletcher, J., and D. Forbes. 2002. An exploration of the persistence of UK unit trusts performance. *Journal of Empirical Finance* 9 (5): 475–493. https://doi.org/10.1016/S0927-5398(02)00006-3.
Frazzini, A., and L. Pedersen. 2013. Betting against beta. *Journal of Financial Economics* 111: 1–25. https://doi.org/10.1016/j.jfineco.2013.10.005.
Frazzini, A., J. Friedman, and L. Pomorski. 2016. Deactivating active share. *Financial Analysts Journal* 72 (2): 14–21. https://doi.org/10.2469/faj.v72.n2.2.
Giglio, S., Liao, Y. & Xiu, D. (2018). Thousands of alpha tests, Chicago Booth Scholl of Business, Working Paper 18-09.
Gregory, A., R. Tharayan, and A. Christidis. 2013. Constructing and testing alternative versions of the Fama-French and Carhart models in the UK. *Journal of Business Finance & Accounting* 40 (1–2): 172–214. https://doi.org/10.1111/jbfa.12006.
Harvey, C.R., and Y. Liu. 2018. Detecting repeatable performance. *Review of Financial Studies* 31 (7): 2499–2552. https://doi.org/10.1093/rfs/hhy014.
Harvey, C.R., and Y. Liu. 2019. Cross-sectional alpha dispersion and performance evaluation. *Journal of Financial Economics* 134: 273–296. https://doi.org/10.1016/j.jfineco.2019.04.005.
Harvey, C.R., Y. Liu, and H. Zhu. 2015. …. and the cross-section of expected returns. *Review of Financial Studies* 29: 5–68. https://doi.org/10.1093/rfs/hhv059.
Hong, H., J.D. Kubik, and J.C. Stein. 2005. Thy neighbor's portfolio: word-of-mouth effects in the holdings and trades of money managers. *The Journal of Finance* 60 (6): 2801–2824. https://doi.org/10.1111/j.1540-6261.2005.00817.x.
Huang, R., K. Pilbeam, and W. Pouliot. 2019. Do actively managed US mutual funds produce positive alpha. *Journal of Economic Behavior and Organization, Forthcoming.*. https://doi.org/10.1016/j.jebo.2019.03.006.
Jiang, H., and M. Verardo. 2018. Does herding behavior reveal skill?. *An Analysis of Mutual Fund Performance, the Journal of Finance* 73 (5): 2229–2269. https://doi.org/10.1111/jofi.12699.
Jones, C., and J. Shanken. 2005. Mutual fund performance with learning across funds. *Journal of Financial Economics* 78: 507–552. https://doi.org/10.1016/j.jfineco.2004.08.009.
Jordan, B.D. & Riley, T.B. (2016). Skill and persistence in mutual fund returns: Evidence from a six-factor model, SSRN.
Jordan, B.D., and T.B. Riley. 2015. Volatility and mutual fund manager skill. *Journal of Financial Economics* 118: 289–298. https://doi.org/10.1016/j.jfineco.2015.06.012.
Kacperczyk, M., and A. Seru. 2007. Fund manager use of public information: New evidence on managerial skills. *The Journal of Finance* 62 (2): 485–528. https://doi.org/10.1111/j.1540-6261.2007.01215.x.
Kacperczyk, M., C. Sialm, and L. Zheng. 2005. On the industry concentration of actively managed mutual funds. *The Journal of Finance* 60 (4): 1983–2011.
Kenchington, D., C. Wan, and H.Z. Yuksel. 2019. Gross profitability and mutual fund performance. *Journal of Banking and Finance* 104: 31–49. https://doi.org/10.1016/j.jbankfin.2019.05.001.
KeswaniStolin, Al.D. 2008. Which money is smart? Mutual fund buys and sells of individual and institutional investors. *The Journal of Finance* 63 (1): 85–118. https://doi.org/10.1111/j.1540-6261.2008.01311.x.
Kosowski, R., A. Timmermann, H. White, and R. Wermers. 2006. Can mutual fund 'stars' really pick stocks? New Evidence from a Bootstrapping Analysis. *The Journal of Finance* 61 (6): 2551–2595. https://doi.org/10.1111/j.1540-6261.2006.01015.x.
Mateus, I.B., C. Mateus, and N. Todorovic. 2016. UK equity mutual fund alphas make a comeback. *International Review of Financial Analysis* 44: 98–110. https://doi.org/10.1016/j.irfa.2016.01.004.
Mateus, I.B., C. Mateus, and N. Todorovic. 2019a. Use of active peer benchmarks in assessing UK mutual fund performance and performance persistence. *The European Journal of Finance* 25 (12): 1077–1098. https://doi.org/10.1080/1351847X.2019.1581639.

Mateus, I.B., C. Mateus, and N. Todorovic. 2019b. Review of new trends in the literature on factor models and mutual fund performance. *International Review of Financial Analysis* 63: 344–354. https://doi.org/10.1016/j.irfa.2018.12.012.

Newey, W., and K. West. 1987. A simple, positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55: 703–708.

Novy-Marx, R. (2014). Understanding defensive equity, working paper, University of Rochester.

Pastor, L., and R.F. Stambaugh. 2002. Mutual fund performance and seemingly unrelated assets. *Journal of Financial Economics* 63: 315–349.

Pastor, L., R.F. Stambaugh, and L.A. Taylor. 2017. Do funds make more when they trade more? *Journal of Finance,* 72: (4), 1483–1528. https://doi.org/10.1111/jofi.12509.

Petajisto, A. 2016. Author response to "Deactivating active share". *Financial Analysts Journal* 72 (4): 11–12. https://doi.org/10.2469/faj.v72.n4.2.

Quigley, G., and R. Sinquefield. 2000. Performance of UK equity unit trusts. *Journal of Asset Management* 1: 72–92.

Sapp, T., and A. Tiwari. 2004. Does stock return momentum explain the "smart money" effect?. *The Journal of Finance*. https://doi.org/10.1111/j.1540-6261.2004.00710.x.

Sensoy, B. 2009. Performance evaluation and self-designated benchmark indexes in the mutual fund industry. *Journal of Financial Economics* 92 (1): 25–39. https://doi.org/10.1016/j.jfineco.2008.02.011.

Sharpe, W.F. 1992. Asset allocation: management style and performance measurement. *Journal of Portfolio Management* 18 (2): 7–19. https://doi.org/10.3905/jpm.1992.409394.

Tonks, I. 2005. Performance persistence of pension fund managers. *Journal of Business* 78: 1917–1942. https://doi.org/10.1086/431447.

Zheng, L. 1999. Is money smart? A study of mutual fund investors' fund selection ability. *The Journal of Finance* 54 (3): 901–933. https://doi.org/10.1111/0022-1082.00131.

Keith Cuthbertson Keith Cuthbertson is Professor of Finance at Bayes Business School, City, University of London. He has worked at HM Treasury, Bank of England, National Institute of Economics and Social Research (NIESR) and at business schools at Imperial College London and the University of Newcastle.

Dirk Nitzsche Dirk Nitzsche is Senior Lecturer in Finance, Course Director of the Quants Masters Programmes and Associate Dean for International Relations at Bayes Business School, City, University of London.

Niall O'Sullivan Niall O'sullivan is Professor of Economics at Cork University Business School, University College Cork. He was formerly an Adjunct Lecturer at Dublin City University Business School, Economist at Bank of Ireland Global Markets, Senior Treasury at FTI Finance and Trading Risk Analyst at Ulster Bank Markets.No error detected.