Forthcoming, *The International Journal of Finance*

Benchmarking UK Mutual Fund Performance: The Random Portfolio Experiment

Andrew Clare[a],

Niall O'Sullivan[b*]

and

Meadhbh Sherman[c]

Abstract
We formally test the age-old question of whether professionally managed equity funds outperform portfolios of stocks selected at random, also known as 'dartboard' or 'monkey' portfolios. We examine the case of UK equity mutual funds between 1980 and 2011. We employ alpha and the t-statistic of alpha as performance measures from CAPM, Fama-French and Carhart factor models. We find that around 5% to 25% of funds across alternative performance measures and models yield abnormal returns beyond that which can be explained by random chance or luck in performance. The t-statistic of alpha indicates a slightly higher percentage of skilful funds compared to alpha, most likely for statistical reasons around short-lived funds. The degree of skilful performance among managed funds is higher when evaluated by a single factor CAPM or Fama and French three factor alpha but a Carhart four factor model explains much of this performance.

JEL classification: G14, G17, C15.
Keywords: mutual fund performance, random portfolios, skill, luck.

[a] The Centre for Asset Management Research, The Sir John Cass Business School, City University, 106 Bunhill Row, EC1Y 8TZ, London, UK. E:a.clare@city.ac.uk. T. +44 207 0405169

[b*] Corresponding Author. Department of Economics and Centre for Investment Research, University College Cork, Ireland. E: niall.osullivan@ucc.ie, T. +353 21 4902765

[c] Department of Economics, University College Cork, Ireland. E: m.sherman@ucc.ie, T. +353 21 4903522

1. Introduction

Ever since Burton Malkiel predicted in *A Random Walk Down Wall Street* that “a blind-folded monkey throwing darts at a newspaper’s financial pages could select a portfolio that would do just as well as one carefully selected by experts”, fund managers have been put to various forms of such a test. The *Wall Street Journal* began an experiment in 1988 run over a decade in which it compared the 'darts' versus the 'pros.' After the 100th dartboard contest the pros came out ahead beating the random selections 61% of the time. However, this then sparked a debate about risk and portfolio size which it was argued biased the outcome. Some say Malkiel’s prediction was too modest and that the monkey would not only match the pros but would beat them.

In a study of the US equity market between 1968 and 2011 Clare, Motson and Thomas (2013) question whether a tracker fund made up of stocks weighted according to their share of the market is an appropriate strategy to build passive investment funds. Ten million simulated portfolios containing stocks randomly selected and randomly weighted from a sample of 1,000 stocks produced higher risk-adjusted returns than a cap-weighted tracker fund over 40 years. One of the suggested implications is that perhaps we should be benchmarking fund managers against such random portfolios rather than against capitalisation weighted indices.

Oddly, the question of random portfolio benchmarking for funds has not received much attention in the academic literature. As suggested by the Capital Asset Pricing Model (CAPM), a capitalization-weighted portfolio should be chosen by investors without stock picking skills for its mean variance efficiency. Arnott et al. (2013) question this. The authors simulate random portfolios for comparison against a capitalisation-weighted benchmark. From 1,000 portfolios, each with 30 equally-weighted stocks from among the 1,000 largest American stocks by market

capitalisation, 99 of the 100 “dartboard portfolios” outperformed a market cap-weighted portfolio of all the 1,000 stocks. A word of caution to investors about the dartboard portfolios, however, is that they embodied high risk (volatility and beta) and tracking error. The central finding in Arnott et al. (2013) is that size and value exposures occur naturally in portfolios. Hence, an equally-weighted dartboard or random portfolio, which over-weights size and value stocks compared to a cap-weighted index benchmark, will yield a premium in terms of raw return or even a CAPM alpha. This should be borne in mind when we evaluate either funds or indices against random portfolios.

The above studies examine the US equity market. Our review of the literature revealed that very little similar work has been undertaken on the UK market. We evaluate the performance of UK equity mutual funds against simulated random, dartboard or monkey portfolios of stocks over the period January 1980 to November 2011. UK mutual fund studies centre on either *ex-post* risk-adjusted performance or *ex-ante* performance persistence (Cuthbertson et al., 2012, 2008; Otten and Reijnders, 2012; Quigley and Sinquefield, 1999; Fletcher, 1997). Risk-adjusted fund performance is most often based on the alpha from the Fama and French (1996) or Carhart (1997) models which control for market, size, value and momentum risk factors. The random portfolio approach here establishes a distribution of performance under the null hypothesis of zero abnormal performance as presumably the monkey selecting stocks has no skill. The performance of actual managed funds can then be compared to this distribution to establish whether it is outside or within that which can be explained by random chance or luck in performance. In short, are funds skilful or could their performance be attributable to luck? In a different approach, Cuthbertson et al. (2008) also examine UK mutual fund performance and distinguish skill from luck in performance by using a non-parametric bootstrap procedure to construct a distribution of random sampling variation in performance or luck. The paper concludes that less than 2% of funds achieve a level of performance beyond that which could

be attributed to chance. Cuthbertson et al. (2012) apply a false discovery rate procedure to UK mutual funds. This method determines the proportion of significant fund alphas that are not just type I errors or 'false discoveries'. The authors find a false discovery rate of around 30% among funds. Cuthbertson et al. (2010) provide a comprehensive survey of both the theory and empirical findings around mutual fund performance globally.

2. Data

Our mutual fund data set is obtained from Morningstar and contains monthly total return data on 1,267 actively managed UK equity unit trusts and Open Ended Investment Companies over the period January 1980 – November 2011. 'UK Equity' funds (by definition) have at least 80% of the fund invested in UK equity. By restricting funds to those investing in UK equity, more accurate performance benchmarks may be used. Our sample includes 658 non-surviving funds. Funds are also categorised by investment styles: 'Equity Income' funds (189 funds) aim to achieve a dividend yield greater than 110% of the market, 'General Equity' funds (901) invest in a broad range of equity while small company funds (177) are invested in stocks which form the lowest 10% of the market by capitalization. The majority of funds (1,050) are onshore UK funds but the sample includes 217 offshore funds, e.g., located in Ireland, Luxembourg, Denmark, France and other European locations. Fund returns are measured before taxes and management fees but capture re-invested dividends (for a like-for-like comparison with our stock returns later).

The number of funds in the sample ranges from 24 in January 1980 growing steadily to 609 funds in November 2011. The number of funds peaks in April 2009 at 629 before falling slightly due to the adverse effects of the financial crisis. The time series average of the monthly

cross-sectional average return is 1.2%, the lowest fund return was -25% in October 1987 while the highest was 12.2% in January 1980. Figure 1 plots the monthly cross-sectional average mutual fund return over the sample period. Average returns of less than 10% occur in October 1987 ('Black Monday'), August 1998 (Russian debt default), September 2001 ('9/11') and in the latter half of 2008 (recent financial crisis). The average monthly cross-sectional standard deviation in returns is 2.05%, leaving scope for the benchmark factors to explain the return variation. Briefly looking at the performance breakdown by investment style, income funds yield an average return of 1.21% with an average standard deviation of 1.34%. Interestingly, general equity funds yield the lowest average return at 1.17% with the highest standard deviation of 1.86%. Small stock funds exhibit the highest average returns of 1.32% with a high average standard deviation of 1.75%.

In our multi-factor performance models the market factor is the FTSE All Share Index of total returns (i.e. including reinvested dividends), taken from Datastream. Excess returns are calculated over the one-month UK T-bill rate, also taken from Datastream. The factor mimicking portfolio for the size effect, small minus big (SMB), is the difference between the monthly returns on the Morgan Stanley Capital International (MSCI) UK small cap index and the returns on the FT 100 index. The value factor, high minus low book-to-market equity (HML), is measured as the difference between the monthly returns of the MSCI UK value index and the returns on the MSCI UK growth index. Both indices are total return measures. These value and growth indices are constructed by MSCI by ranking all the stocks in their UK national index by their book-to-market ratio. Starting with the highest book-to-market ratio stocks, these are assigned to the value index until 50% of the market capitalization of the national index is reached. The remaining stocks are assigned to the growth index. The MSCI national indices have a market coverage of at least 85%. The small cap, value and growth indices were obtained from MSCI at

www.msci.com. The FT 100 index is taken from Datastream. We construct the mimicking portfolio for momentum, MOM, using the historic constituents of the FTSE All Share index. Each month the constituent stocks are ranked by their past 11 month return, portfolios of the top and bottom 30% of stocks are then held for one month. The momentum benchmark portfolio is formed as the difference in the holding period return between the top and bottom portfolios.

We conduct our analysis on the historic constituents of the FTSE All Share index, i.e., we cross-reference with the London Share Price Database (LSPD) Archive file which records the constituents of the FTSE All Share back through time. This leaves us with the comprehensive universe of liquid stocks that UK equity mutual funds realistically chose from. Both our fund return and stock return measures are total returns and include re-invested dividends.

3. Methodology

The purpose of our paper is to compare the performance of actual UK equity mutual funds against a distribution of performance based on chance or luck as measured by the performance of portfolios of stocks selected at random. We examine equity funds so that more accurate performance benchmarks may be constructed. In particular, we measure performance as the risk-adjusted alpha from three alternative well-established factor models below including the CAPM [1], the Fama and French (1996) three factor model [2] and the Carhart (1997) four factor model [3], which also controls for momentum in performance attribution.

$$r_{i,t} = \alpha_i + \beta_i(r_{m,t}) + \varepsilon_{1,t} \qquad [1]$$

$$r_{i,t} = \alpha_i + \beta_i(r_{m,t}) + \beta_2(SMB_t) + \beta_3(HML_t) + \varepsilon_{2,t} \quad [2]$$

$$r_{i,t} = \alpha_i + \beta_i(r_{m,t}) + \beta_2(SMB_t) + \beta_3(HML_t) + \beta_4(MOM_t) + \varepsilon_{3,t} \quad [3]$$

where $r_{i,t}$ is the excess return (over the one month UK T-Bill rate) on fund *i* at time *t*, $r_{m,t}$ is the excess return on the market factor at time *t*, SMB_t, HML_t and MOM_t are the size, value and momentum risk factors at time *t*.

We first construct a distribution of alpha which exhibits only random chance or luck around a true value which represents no abnormal performance. We then measure where the alpha of actual managed mutual funds lies within this distribution. We calculate p-values which indicate the probability of observing the actual alpha of a given fund by chance only. Specifically, for each fund *i* in our sample, *i* = 1, 2…1,267, each month we construct a portfolio of stocks of portfolio size *s* = 30, (similar to Arnott et al. 2013), over the same time period as fund *i* exists. Each month the *s* stocks are selected randomly from among the constituent stocks of the FTSE All Share index that month and the equally-weighted portfolio return is calculated[1,2]. This generates a monthly time series of portfolio raw returns over the same time period as fund *i* exists. We then estimate the CAPM, Fama and French and Carhart alphas of this simulated random portfolio[3]. We denote this alpha as $\tilde{\alpha}_i$. Because the portfolio is generated randomly, it is not expected to exhibit any selectivity skill. We repeat this simulation process 1,000 times to

[1] The London Share Price Database (LSPD) Archive file provides information on the monthly historic constituents of the FTSE All Share index. The number of constituents varies through time but is approx. 1,000 stocks.

[2] We also constructed value weighted portfolios using capitalisation data from the LSPD but the overall results were qualitatively very similar. We also experimented with value of *s* ranging from 25 to 50 but this did not alter our conclusions.

[3] The reader may wonder why we adjust the random portfolios for size, value and momentum risk since if the stocks are selected randomly the portfolios should not be exposed to such risks. However, Arnott et al (2013) find that even random portfolios introduce size and value "tilt". An equal weighted random portfolio will over-weight small and value stocks and 'benefit' from a premium compared to a cap-weighted benchmark.

generate a distribution of $\tilde{\alpha}_{i,j}$, j = 1, 2…1,000. The simulation procedure generates both positive and negative values of $\tilde{\alpha}_{i,j}$, the scale of which provide an indication of how well or badly a portfolio of stocks could perform in terms of alpha by random chance only. We then compare the actual alpha of fund *i*, denoted $\hat{\alpha}_i$ against the distribution of $\tilde{\alpha}_{i,j} = [\tilde{\alpha}_{i,1}, \tilde{\alpha}_{i,2}, \ldots . \tilde{\alpha}_{i,1,000}]$. The p-value of fund *i* is the percentage of the 1,000 $\tilde{\alpha}_{i,j}$ values which are greater than $\hat{\alpha}_i$. For example, suppose only 50 of the $\tilde{\alpha}_{i,j}$ values are greater than $\hat{\alpha}_i$, this implies that the p-value of fund *i* is 0.05, or that there is only a 5% chance of observing fund *i*'s actual alpha due to random chance or luck.

In order to improve the statistical reliability of our findings, we restrict our analysis to funds with at least 36 monthly observations. Furthermore, to test the robustness of our findings, we also repeat the above simulation procedure using the t-statistic of alpha, rather than alpha, as our performance measure since it has superior statistical properties, Cuthbertson et al. (2008). Short-lived funds with few observations are likely to have alpha estimates with a high standard error. In turn, this creates large variability in the distribution of the 1,000 alphas from the randomly generated portfolios, leading to longer tails in the distribution and in turn leading to higher p-value estimates. The t-statistic of alpha which incorporates the standard error helps mitigate this issue. In addition, the t-statistic helps mitigate survival bias problems (Brown, Goetzmann, Ibbotson and Ross (1992)). All of our t-statistics are Newey-West adjusted for a lag order of 2, which we found to be the most appropriate lag order.

While Cuthbertson et al. (2008) also distinguish luck from skill in performance, our methodological approach is quite different in this paper. The former bootstrap residuals from the

OLS estimation of a performance model for each fund (several alternative performance models are studied) and simulate a time series of performance under the null hypothesis of alpha equal to zero. They then re-estimate the performance model using the simulated returns and estimate alpha, which then represents random sampling variation in alpha around a true value equal to zero. By carrying out 1,000 bootstrap simulations they are able to generate a nonparametric distribution of alpha under the null hypothesis where its true value equals zero. This distribution represents the level of alpha performance that could be attributable to random sampling variation, chance or luck. They then compare actual fund alphas against this distribution and calculate performance p-values. They perform the same procedure employing the t-statistic of alpha as the performance measure[4]. The nonparametric procedure is a key strength of the paper as it mitigates against non-normality in the model residuals, which is found to be quite prevalent. In our paper, we construct a luck based distribution of alpha quite differently as an alternative: as our simulated portfolios are generated randomly, they are not based on any skill and the resulting distribution of alpha is attributable to luck only. We then compare actual or real fund alphas against this distribution.

4. Empirical Results

The results of our analysis are presented in Tables 1 and 2. Table 1 reports findings where fund performance is based on alpha while Table 2 reports results based on the t-statistic of alpha as the performance measure. Panels A, B and C in Table 1 report results around alpha from the three alternative performance models. Panel A refers to alpha findings based on the single factor CAPM, Panel B presents alpha findings based on the Fama and French three factor

[4] The reader is referred to Cuthbertson et al. (2008) for a complete discussion of the procedure. In each bootstrap simulation the procedure sorts the simulated alphas from highest to lowest. This generates 1,000 highest simulated alphas under the null against which the top actual fund alpha is compared, 1,000 second highest alphas under the null against which the second highest alpha is compared and so on.

model, while Panel C reports alpha findings based on the Carhart four factor model. After restricting the analysis to funds with at least 36 monthly observations there are 652 funds in the sample. In each panel the row denoted “Alpha” shows actual mutual fund alphas at various points in the cross-sectional distribution ranging from the lowest alpha, denoted “min”, to the highest alpha, denoted “max”. When we use the single factor CAPM the highest alpha is 1.467% per month. The row denoted “Average” refers to the average of the 1,000 alphas of the randomly generated portfolios corresponding to the fund in that column. The row headed “p-value” refers to the p-value of the actual alpha in row 1. Again, according to the single factor CAPM in Panel A, the best fund has a p-value of 0.00. This indicates that none of the 1,000 alphas from the randomly generated portfolios over the same time period are greater than 1.467. The average of these 1,000 alphas is 0.712. “Obs” refers to the number of monthly observations of the fund. “Style” denotes whether the fund is an income fund (1), a general equity fund (2) or a small stock fund (3). “On/Off” indicates whether the fund is an onshore fund (1) or an offshore fund (0). Finally, “Survivor” denotes whether the fund is a survivor (1) or a non-survivor (0). The highest alpha fund has existed for 36 months and is an onshore surviving small stock fund. From Panel B, the highest Fama and French three factor alpha is 0.921% per month (a different fund from the best fund in Panel A) and, again, has a p-value of 0.00. In Panel C, the highest Carhart four factor alpha is 1.059% per month with a p-value of 0.00. Therefore, across all three factor models the fund with the best alpha has a p-value of 0.00 indicating that the managers of these funds have demonstrated stock selectivity skill (as inferred by alpha) greater than that attributable to chance. If we work with the conventional 5% statistical significance level, this means that a p-value up to 0.05 is required for us to draw the same conclusion.

From Table 1, using the single factor model in Panel A, all funds ranked down to 'max 5%', i.e., the 95th percentile, we categorise as having skill since their p-values are less than 0.05. With 652 funds in the sample these are the top 33 funds by alpha. Indeed, in results not shown, when we focus on the distribution more closely between the 95th and 90th percentiles, we find that the top 38 funds by alpha demonstrate skill. Of course, there is no direct link between the level of alpha and whether the fund manager is skilful or not. For example, the fund ranked 39th has an alpha of 0.566% per month with a p-value of 0.095 (not skilful at 5% significance) but the fund ranked 41st has an alpha of 0.563% with a p-value of 0.00. In total, 89 funds, or 14% of the sample, have p-values less than 0.05 and may be deemed skilful, though they are not necessarily the top 89 funds as ranked by alpha. From Panel B, where results are based upon the three factor model which controls for fund performance attributable to size and value risk factors, skill is found in all funds ranked in the top 10% by alpha. In fact, again in results not shown, a total of 19% of the sample (124 funds) exhibit skill at the 5% significance level (though, again, not necessarily the top 124 funds by alpha). This greater degree of skilful performance among actual funds found using the Fama and French three factor model compared to the single factor CAPM is due to the fact that, as found by Arnott et al. (2013), the random portfolios exhibit a size tilt in many cases. Controlling for this reduces the size risk-adjusted alpha in the random portfolios making it 'easier' for actual funds to outperform the random portfolios. In Panel C in the case of the Carhart model, only funds ranked down to 20th position by alpha demonstrate significant p-values, or 34 funds in total have p-values less than 0.05. This reduced prevalence of skilful funds according to the four factor model compared to the three factor model suggests that much of the performance of actual mutual funds is attributable to momentum: the four factor alpha explains much of the observed CAPM and three factor alpha in actual funds.

Looking at the characteristics of the top performing skilful funds, it is noteworthy from all three panels of Table 1 that shorter-lived funds are slightly more prevalent, especially in the extreme tail of the distribution, the funds are onshore, surviving funds and small stock funds are disproportionately represented even after controlling for a size risk factor in the models in Panels B and C.

The poor performing end of the performance distribution is also revealing. The results based on the single factor CAPM in Panel A show that the worst performing fund has an alpha of -1.06% per month over its 52 months in existence. A p-value of 0.00 indicates that none of the 1,000 alphas from the randomly generated portfolios over the same time period are less than -1.06%, i.e., there is a 0% probability of observing an alpha of -1.06% by random chance only over the same time period: the poor performance of this fund is outside the boundary of that which can be explained by bad luck. The same interpretation applies to funds with p-values of 0.05 or less in the left tail of the cross-sectional distribution. A similar conclusion may be drawn about performance at the lower end of the distribution across all three performance models.

Table 2 provides similar information except that the performance measure used to evaluate funds is the t-statistic of alpha, which has superior statistical properties particularly in the evaluation of short lived funds. When using the single factor CAPM (Panel A) the best performing fund has a t-statistic of alpha of 4.171. The corresponding p-value of zero indicates that none of the t-statistics of the 1,000 alphas from the randomly generated portfolios over the same time period are greater than 4.171, i.e., there is a zero probability of observing this level of performance by chance: we may deem the returns on this fund to have been generated with

some skill. The statistics in Panels B and C of Table 2 have a similar interpretation. Comparing the p-values of the funds in the upper end of the performance distributions in panels A, B and C of Table 2 with those of Table 1, it can be seen that when we analyse by the t-statistic of alpha in Table 2 a higher percentage of the sample of funds are skilful at the 5% significance level. Specifically, when we use the single factor model, a total of 22% of the funds are skilful (though this is not necessarily the top 22% as ranked by the t-statistic of alpha). This compares to 14% when judged using alpha. According to the Fama and French three factor model 24% of funds are skilful, compared to 19% when judged using alpha while, according to the Carhart four factor model, 8% of funds are skilful, compared to 5% when judged using alpha. Therefore, according to the t-statistic of alpha, which better mitigates against noise in the performance estimate in particular for short lived funds, we find a higher degree of skill among funds. Consistent with this, the results in Table 2 reveal that the high performing funds tend to be longer lived in contrast to the findings in Table 1. The results in Table 2 also reveal that among the high performing funds, general equity funds tend to dominate with some income funds also represented. However, small stocks funds are far less well represented in contrast to the findings in Table 1.

4.1 Robustness Tests

The role of momentum in fund performance is well established and documented, see, for example, Carhart (1997). In constructing the benchmark momentum risk factor, however, one must select the length of the backward looking time window over which the universe of stock are ranked. In this study we employ an 11 month backward looking ranking period length which is common in the fund performance and asset pricing literature (Cuthbertson et al., 2008; Foran and O'Sullivan, 2014; Foran et al., 2014). However, the selection could have a significant bearing on the momentum factor and the evaluation of funds. In order to examine the robustness of our findings around this question, we repeat our analysis where in the Carhart

four factor model we construct a momentum risk factor using both a six month and three month backward looking time period when ranking stocks. We report the results from these tests in Table 3, where Panel A and Panel B relate to the six month and three month momentum factors respectively. To conserve space we limit our results to those for the t-statistic of alpha as the performance measure. Comparing the findings of Panes A and B of Table 3 with those of Panel C in Table 2, it is clear our initial conclusions are unaltered. The level of outperformance among funds in the right tail of the distribution, as indicated by a p-value < 0.05, is the same across all three panels - our findings are robust to the choice of momentum factor in the model.

In a second test of the robustness of our findings it may be interesting to examine the performance of funds in periods of tranquillity versus crisis. Over our long sample period between January 1980 and November 2011 there are, of course, many such periods of tranquillity versus crisis one could examine. However, one obvious candidate period of interest is the most recent financial crisis period. We examine whether the exclusion of this period would have a significant impact on our findings. We take the financial crisis period as post August 2007. (Specifically, on the 9th August 2007 Bloomberg reported that BNP Paribas halted withdrawals from three investment funds because it couldn't "fairly" value their holdings after U.S. subprime mortgage losses roiled credit markets. We use this as one of the early indications of the financial crisis). In Table 4, we present our skill versus luck findings over two sample periods. Again, to conserve space we limit our results to those for the t-statistic of alpha as the performance measure. First, in Panel A, we present the results for the full period January 1980 – November 2011 as before. In Panel B, we present results over the shorter period January 1980 – July 2007 inclusive, i.e., excluding the financial crisis period. Comparing the results between Panel A and B, we find no difference in the interpretation of results. The degree of performance not attributable to chance is qualitatively unchanged.

Overall, we find evidence of abnormal returns in a small proportion of UK mutual funds which cannot be explained by random chance or luck in performance. Employing the t-statistic

of alpha as the performance measure points to a slightly higher percentage of skilful funds compared to the findings based on alpha. This is most likely because, by dividing by the standard error, the distribution of the t-statistics of the alphas of the randomly generated portfolios have shorter tails which increases the likelihood of finding lower p-values. The degree of skilful performance is higher when evaluated by a single factor CAPM or the Fama and French three factor model. In the Carhart four factor model when mutual returns attributable to momentum effects are discounted the prevalence of skilful performance is reduced.

5. Conclusion

We test whether the stock selection performance of UK equity mutual funds, as measured by both alpha as well the t-statistic of alpha, is greater than that which can be explained by random chance or luck as measured by the distribution of the performance of randomly generated portfolios of stocks under the null hypothesis of no abnormal performance. We find that around 5% to 25% of funds across alternative performance measures and models yield abnormal returns beyond that which can be explained by random chance or luck in performance, i.e., the managers responsible for generating these returns may be considered to have possessed some skill. Employing the t-statistic of alpha as the performance measure points to a slightly higher percentage of skilful funds compared to alpha, most likely for statistical reasons around short-lived funds. The degree of skilful performance is higher when evaluated using a single factor CAPM or Fama and French three factor model but is reduced by the Carhart four factor model when actual mutual returns attributable to momentum effects are discounted. The high performing funds tend to be general equity funds, rather than income funds or small stock funds, they tend to be surviving funds and are generally managed from onshore UK rather than offshore locations.

References

Arnott, R., Hsu, J., Kalesnik, V. and Tindall, P., (2013), The Surprising Alpha From Malkiel's Monkey and Upside-Down Strategies, *Journal of Portfolio Management*, Summer 2013, vol. 39, no. 4, pp. 91-105.

Brown, S., Goetzmann, W., Ibbotson, R. and Ross, S., (1992), "Survivorship Bias in Performance Studies", *Review of Financial Studies,* vol*.* 5, issue 4, pp. 553-580.

Carhart, M., (1997), On persistence in mutual fund performance, *Journal of Finance*, vol. 52, no. 1, pp. 57-82.

Clare, A., Motson, N. and Thomas, S., (2013), An evaluation of alternative equity indices Part 1: Heuristic and optimised weighting schemes, viewed 23 October 2014, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2242028.

Cuthbertson, K., Nitzsche, D. and O'Sullivan, N., (2008), UK Mutual Fund Performance: skill or luck? *Journal of Empirical Finance*, vol. 15, issue 4, pp. 613-634.

Cuthbertson, K., Nitzsche, D. and O' Sullivan, N., (2010), Mutual Fund Performance: Measurement and Evidence, *Journal of Financial Markets, Instruments and Institutions*, vol. 19, issue 2, pp. 95-187.

Cuthbertson, K., Nitzsche, D. and O' Sullivan, N., (2012), False Discoveries in UK Mutual Fund Performance, *European Financial Management*, vol. 18, issue 3, pp. 444-463.

Fama, E. and French, K., (1996), Multifactor explanations of asset pricing anomalies, *Journal of Finance*, vol. 51, issue 1, pp. 55-84.

Fletcher, J., (1997), An examination of UK unit trust performance within the arbitrage pricing framework, *Review of Quantitative Finance and Accounting,* vol. 8, issue 2, pp. 91-107.

Foran, J. and O'Sullivan, N., (2014), Liquidity Risk and the Performance of UK Mutual Funds, *International Review of Financial Analysis*, 35, 178–189.

Foran, J., Hutchinson, M. and O'Sullivan, N., (2014), The asset pricing effects of UK market liquidity shocks: Evidence from tick data, *International Review of Financial Analysis*, 32, 85–94.

Malkiel, B., (1973), *A Random Walk Down Wall Street: The Time-Tested Strategy for Successful Investing*. W.W. Norton, New York.

Newey, W. and West, K., (1987), 'A Simple, Positive Semi-Definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix', *Econometrica*, vol. 55, no. 3, pp. 703-708.

Otten, R. and Reijnders, M., (2012), The Performance of Small Cap Mutual Funds: Evidence for the UK, viewed on 23 October 2014, http://ssrn.com/abstract=2056932

Quigley, G. and Sinquefield, R., (1999), The performance of UK equity unit trusts, Report by Dimensional Fund Advisors for Institute for Fiduciary Education.

Figure 1: UK Mutual Fund Returns 1980 – 2011
The chart plots the monthly time series of the cross-sectional average mutual return from January 1980 to November 2011.

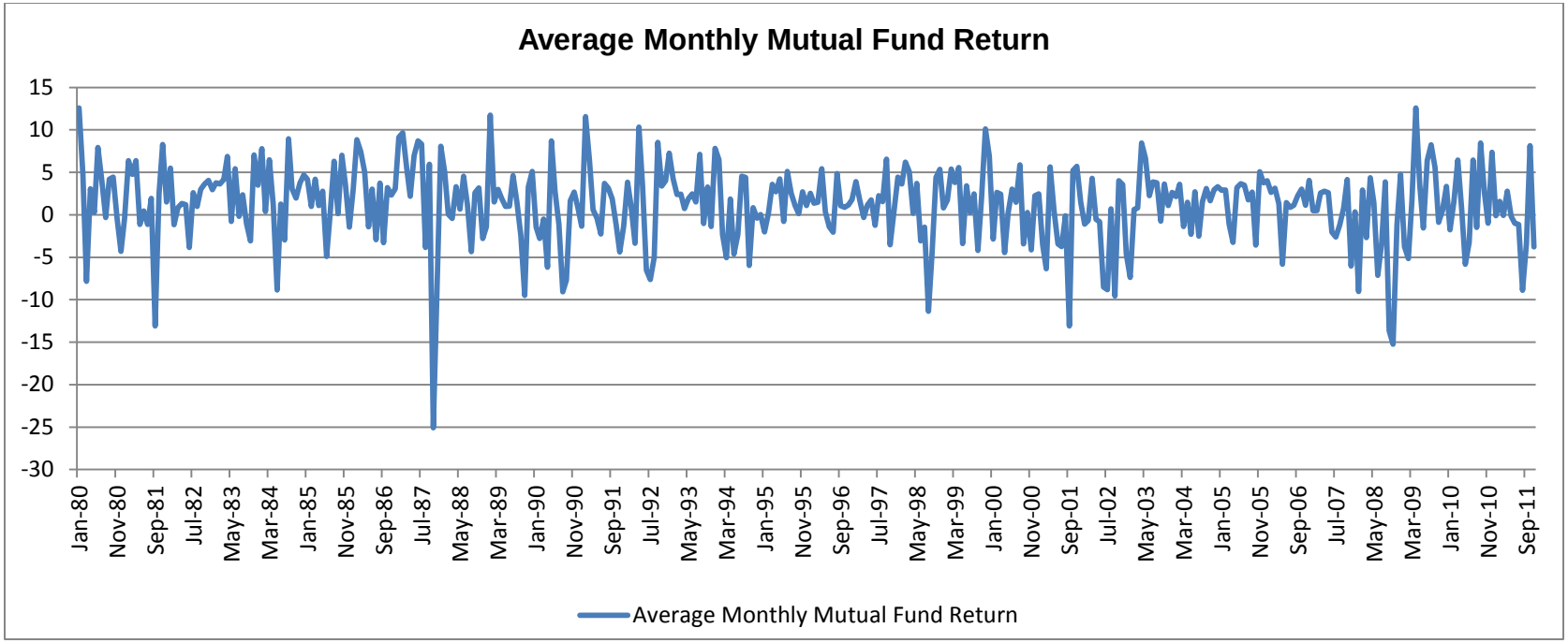

Table 1: Mutual Fund Performance Evaluation Against Randomly Generated Alpha

The row denoted "Alpha" shows mutual fund actual alphas at various points in the cross-sectional distribution, e.g., "min" corresponds to the lowest alpha, "5 min" corresponds to the fifth lowest alpha, "min5%" corresponds to the alpha at the fifth percentile, "max" corresponds to the highest alpha etc. "Average" refers to the average of the 1,000 alphas of the randomly generated portfolios corresponding to the fund in that column. "p-value" refers to the p-value of the actual alpha in row 1. "Obs" refers to the number of monthly observations of the fund. "Style" denotes whether the fund is an income fund (1), a general equity fund (2) or a small stock fund (3). "On/Off" refers to whether the fund is an onshore fund (1) or an offshore fund (0). "Survivor" denotes whether the fund is a survivor (1) or a non-survivor (0). Panels A, B and C report results around alpha from the three alternative performance models as indicated.

Panel A: Single Factor CAPM

$$(R_i - r_f)_t = \alpha_i + \beta_{1i}(R_m - r_f)_t + \varepsilon_{it}$$

	min	5 min	min5%	min10%	min40%	max30%	max10%	max5%	20max	15max	12max	10max	7 max	5 max	3 max	2 max	max
Alpha	-1.06	-0.698	-0.264	-0.171	0.040	0.203	0.458	0.583	0.708	0.751	0.818	0.845	1.007	1.050	1.192	1.459	1.467
Average	0.016	-0.018	0.126	0.027	0.038	0.031	0.068	0.285	0.284	0.122	0.233	0.088	0.036	0.088	0.235	0.678	0.712
p-value	0	0.005	0.04	0.23	0.55	0.27	0.055	0.04	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
Obs.	52	55	71	66	214	52	65	122	152	178	143	74	66	196	163	36	36
Style	2	2	2	2	1	3	3	2	2	3	3	2	3	3	3	3	3
On/Off	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Survivor	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Panel B: Fama and French Three Factor Model

$$(R_i - r_f)_t = \alpha_i + \beta_{1i}(R_m - r_f)_t + \beta_{2i}SMB_t + \beta_{3i}\,HML_t + \varepsilon_{it}$$

	min	5 min	min5%	min10%	min40%	max30%	max10%	max5%	20max	15max	12max	10max	7 max	5 max	3 max	2 max	max
Alpha	-0.99	-0.664	-0.340	-0.222	-0.017	0.135	0.635	0.417	0.534	0.631	0.633	0.741	0.808	0.889	0.901	0.913	0.921
Average	0057	0.050	0.154	0.164	-0.107	0.06	-0.060	-0.053	-0.037	-0.031	0.064	-0.048	-0.032	0.140	0.145	-0.034	0.034
p-value	0	0.005	0.05	0.15	0.730	0.42	0.00	0.01	0.00	0.00	0.025	0.00	0.00	0.005	0.00	0.00	0.00
Obs.	92	55	36	43	191	60	138	152	151	79	59	130	163	36	36	45	66
Style	2	2	1	2	1	1	1	2	1	2	2	3	3	3	3	3	3
On/Off	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Survivor	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1

Panel C: Carhart Four Factor Model

$$(R_i - r_f)_t = \alpha_i + \beta_{1i}(R_m - r_f)_t + \beta_{2i}SMB_t + \beta_{3i}\,HML_t + \beta_{4i}\,MOM_t + \varepsilon_{it}$$

	min	5 min	min5%	min10%	min40%	max30%	max10%	max5%	20max	15max	12max	10max	7 max	5 max	3 max	2 max	max
Alpha	-0.98	-0.661	-0.339	-0.240	-0.027	0.128	0.338	0.407	0.523	0.619	0.633	0.678	0.757	0.892	0.917	0.934	1.059
Average	0.204	0.139	0.037	0.143	0.157	0.123	0.113	0.165	0.153	0.132	0.154	0.181	0.097	0.148	0.002	0.053	0.153
p-value	0.00	0.02	0.00	0.035	0.21	0.48	0.08	0.10	0.04	0.025	0.005	0.015	0.01	0.015	0.005	0.00	0.00
Obs.	92	55	346	110	124	113	148	158	143	79	163	130	45	36	45	45	66
Style	2	2	3	2	2	2	2	2	3	2	3	3	3	3	3	3	3
On/Off	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Survivor	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Table 2: Mutual Fund Performance Evaluation Against Randomly Generated t-statistics of Alpha

The row denoted "t-Alpha" shows mutual fund t-statistics of alpha at various points in the cross-sectional distribution, e.g., "min" corresponds to the lowest t-statistic of alpha, "5 min" corresponds to the fifth lowest t-alpha, "min5%" corresponds to the t-alpha at the fifth percentile, "max" corresponds to the highest t-alpha etc. "Average" refers to the average of the 1,000 t-statistics of alpha of the randomly generated portfolios corresponding to the fund in that column. "p-value" refers to the p-value of the actual t-statistic of alpha in row 1. "Obs" refers to the number of monthly observations of the fund. "Style" denotes whether the fund is an income fund (1), a general equity fund (2) or a small stock fund (3). "On/Off" refers to whether the fund is an onshore fund (1) or an offshore fund (0). "Survivor" denotes whether the fund is a survivor (1) or a non-survivor (0). Panels A, B and C report results around the t-statistics of alpha from the three alternative performance models as indicated.

Panel A: Single Factor CAPM

$$(R_i - r_f)_t = \alpha_i + \beta_{1i}(R_m - r_f)_t + \varepsilon_{it}$$

	min	5 min	min5%	min10%	min40%	max30%	max10%	max5%	20max	15max	12max	10max	7 max	5 max	3 max	2 max	max
t-Alpha	-7.08	-2.926	-1.486	-1.044	0.284	1.092	2.004	2.501	2.772	2.893	2.982	3.068	3.202	3.225	3.286	3.436	4.171
Average	0.189	0.437	0.827	0.111	0.751	0.059	0.132	1.024	1.022	0.762	0.928	0.942	0.892	0.500	0.844	0.788	0.577
p-value	0.00	0.00	0.00	0.02	0.21	0.04	0.01	0.01	0.005	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Obs.	43	60	110	71	105	66	337	152	151	138	122	122	148	101	148	104	130
Style	2	2	1	2	3	3	2	2	1	1	2	2	2	2	2	2	3
On/Off	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Survivor	0	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1

Panel B: Fama and French Three Factor Model

$$(R_i - r_f)_t = \alpha_i + \beta_{1i}(R_m - r_f)_t + \beta_{2i}SMB_t + \beta_{3i}\,HML_t + \varepsilon_{it}$$

	min	5 min	min5%	min10%	min40%	max30%	max10%	max5%	20max	15max	12max	10max	7 max	5 max	3 max	2 max	max
t-Alpha	-7.60	-3.220	-2.022	-1.501	-0.105	0.859	1.829	2.434	2.688	2.832	2.908	2.985	3.045	3.137	3.162	3.178	5.172
Average	0.353	-0.914	-1.304	-0.016	-0.46	0.187	0.032	-0.249	0.370	-0.324	0.064	0.125	0.060	0.296	0.068	-0.042	-0.249
p-value	0.00	0.005	0.20	0.04	0.64	0.23	0.02	0.00	0.015	0.00	0.00	0.00	0.00	0.005	0.00	0.00	0.00
Obs.	43	141	346	92	182	56	84	140	105	178	68	122	117	104	65	119	130
Style	2	2	2	2	3	2	1	2	1	3	2	2	2	2	2	2	3
On/Off	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1
Survivor	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Panel C: Carhart Four Factor Model

$$(R_i - r_f)_t = \alpha_i + \beta_{1i}(R_m - r_f)_t + \beta_{2i}SMB_t + \beta_{3i}\,HML_t + \beta_{4i}\,MOM_t + \varepsilon_{it}$$

	min	5 min	min5%	min10%	min40%	max30%	max10%	max5%	20max	15max	12max	10max	7 max	5 max	3 max	2 max	max
t-Alpha	-7.88	-3.202	-2.144	-1.625	-0.151	0.875	1.949	2.459	2.701	2.909	2.924	2.974	3.029	3.137	3.262	3.803	4.720
Average	0.447	0.225	0.359	0.548	0.501	0.810	0.466	0.708	0.558	0.631	0.581	0.455	0.551	0.918	0.660	0.684	0.796
p-value	0.00	0.00	0.00	0.04	0.25	0.47	0.06	0.015	0.03	0.03	0.00	0.015	0.00	0.015	0.00	0.00	0.00
Obs.	43	372	55	53	37	66	36	178	196	68	117	59	249	104	277	277	130
Style	2	2	2	2	2	2	3	3	3	2	2	2	1	2	2	1	3
On/Off	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Table 3: Mutual Fund Performance Evaluation Against Randomly Generated t-statistics of Alpha for alternative momentum measures
The row denoted "t-Alpha" shows mutual fund t-statistics of alpha at various points in the cross-sectional distribution, e.g., e.g., "min" corresponds to the lowest t-statistic of alpha, "5 min" corresponds to the fifth lowest t-alpha, "min5%" corresponds to the t-alpha at the fifth percentile, "max" corresponds to the highest t-alpha etc. "Average" refers to the average of the 1,000 t-statistics of alpha of the randomly generated portfolios corresponding to the fund in that column. "p-value" refers to the p-value of the actual t-statistic of alpha in row 1. "Obs" refers to the number of monthly observations of the fund. "Style" denotes whether the fund is an income fund (1), a general equity fund (2) or a small stock fund (3). "On/Off" refers to whether the fund is an onshore fund (1) or an offshore fund (0). "Survivor" denotes whether the fund is a survivor (1) or a non-survivor (0). Panel A and Panel B report results for the Carhart four factor model where the momentum factor is generated looking back 6 months and 3 months respectively with a holding period of one month in both cases.

Panel A: Carhart Four Factor Model (6 month momentum)

$$(R_i - r_f)_t = \alpha_i + \beta_{1i}(R_m - r_f)_t + \beta_{2i}SMB_t + \beta_{3i}\,HML_t + \beta_{4i}\,MOM_t + \varepsilon_{it}$$

	min	5 min	min5%	min10%	min40%	max30%	max10%	max5%	20max	15max	12max	10max	7 max	5 max	3 max	2 max	max
t-Alpha	-7.72	-3.168	-2.037	-1.646	-0.163	0.835	1.836	2.391	2.692	2.815	2.838	2.877	3.00	3.079	3.166	3.693	4.722
Average	0.837	0.026	0.945	0.613	0.728	0.648	0.752	0.296	0.801	0.751	0.629	0.720	0.657	1.042	0.985	0.571	0.775
p-value	0.00	0.00	0.00	0.02	0.15	0.42	0.14	0.01	0.02	0.01	0.00	0.02	0.00	0.02	0.02	0.00	0.00
Obs.	43	372	90	222	181	171	112	45	151	122	117	101	119	104	65	277	130
Style	2	2	2	3	2	1	2	3	1	2	2	2	2	2	2	1	3
On/Off	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1
Survivor	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Panel B: Carhart Four Factor Model (3 month momentum)

$$(R_i - r_f)_t = \alpha_i + \beta_{1i}(R_m - r_f)_t + \beta_{2i}SMB_t + \beta_{3i}\,HML_t + \beta_{4i}\,MOM_t + \varepsilon_{it}$$

	min	5 min	min5%	min10%	min40%	max30%	max10%	max5%	20max	15max	12max	10max	7 max	5 max	3 max	2 max	max
t-Alpha	-7.62	-3.104	-1.938	-1.494	-0.119	0.912	1.941	2.445	2.755	2.884	2.935	2.946	3.006	3.049	3.209	3.713	5.078
Average	0.853	0.580	0.108	0.610	0.587	0.439	0.459	-0.066	0.363	0.770	-0.042	0.327	0.427	0.282	0.692	0.036	0.373
p-value	0.00	0.00	0.01	0.00	0.19	0.26	0.16	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Obs.	43	61	120	71	94	117	133	280	148	105	249	151	122	119	65	277	130
Style	2	2	2	2	2	2	2	3	2	1	1	1	2	2	2	1	3
On/Off	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1
Survivor	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1

Table 4: Mutual Fund Performance Evaluation Against Randomly Generated t-statistic of Alpha – Excluding the Financial Crisis Period
The row denoted "t-Alpha" shows mutual fund t-statistics of alpha at various points in the cross-sectional distribution, e.g., e.g., "min" corresponds to the lowest t-statistic of alpha, "5 min" corresponds to the fifth lowest t-alpha, "min5%" corresponds to the t-alpha at the fifth percentile, "max" corresponds to the highest t-alpha etc. "Average" refers to the average of the 1,000 t-statistics of alpha of the randomly generated portfolios corresponding to the fund in that column. "p-value" refers to the p-value of the actual t-statistic of alpha in row 1. "Obs" refers to the number of monthly observations of the fund. "Style" denotes whether the fund is an income fund (1), a general equity fund (2) or a small stock fund (3). "On/Off" refers to whether the fund is an onshore fund (1) or an offshore fund (0). "Survivor" denotes whether the fund is a survivor (1) or a non-survivor (0). Panel A and Panel B report results for the Carhart four factor model over the sample period as indicated.

Panel A: Carhart Four Factor Model: January 1980 – December 2011

$$(R_i - r_f)_t = \alpha_i + \beta_{1i}(R_m - r_f)_t + \beta_{2i}SMB_t + \beta_{3i}\,HML_t + \beta_{4i}\,MOM_t + \varepsilon_{it}$$

	min	5 min	min5%	min10%	min40%	max30%	max10%	max5%	20max	15max	12max	10max	7 max	5 max	3 max	2 max	max
t-Alpha	-7.08	-2.926	-1.486	-1.044	0.284	1.092	2.004	2.501	2.772	2.893	2.982	3.068	3.202	3.225	3.286	3.436	4.171
Average	0.189	0.437	0.827	0.111	0.751	0.059	0.132	1.024	1.022	0.762	0.928	0.942	0.892	0.500	0.844	0.788	0.577
p-value	0.00	0.00	0.00	0.02	0.21	0.04	0.01	0.01	0.005	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Obs.	43	60	110	71	105	66	337	152	151	138	122	122	148	101	148	104	130
Style	2	2	1	2	3	3	2	2	1	1	2	2	2	2	2	2	3
On/Off	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Survivor	0	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1

Panel B: Carhart Four Factor Model: January 1980 – July 2007

$$(R_i - r_f)_t = \alpha_i + \beta_{1i}(R_m - r_f)_t + \beta_{2i}SMB_t + \beta_{3i}\,HML_t + \beta_{4i}\,MOM_t + \varepsilon_{it}$$

	min	5 min	min5%	min10%	min40%	max30%	max10%	max5%	20max	15max	12max	10max	7 max	5 max	3 max	2 max	max
t-Alpha	-5.51	-2.966	-1.884	-1.437	-0.089	0.998	2.096	2.524	2.620	2.758	3.012	3.054	3.350	3.755	4.190	4.266	5.86
Average	-0.12	0.055	-0.180	-0.186	0.081	-1.038	-0.285	-0.982	0.289	0.301	0.247	-0.134	-0.136	-0.295	-0.009	-0.011	0.231
p-value	0.00	0.00	0.04	0.11	0.430	0.01	0.005	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Obs.	72	59	69	67	56	311	229	321	100	93	101	66	66	226	151	71	79
Style	2	2	2	2	2	2	3	1	1	2	2	2	2	1	3	2	1
On/Off	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	3
Survivor	1	1	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1