



City Research Online

City St George's, University of London

Citation: Huang, R., Pilbeam, K. & Pouliot, W (2021). Do Actively Managed US Mutual Funds Produce Positive Alpha?. *Journal of Economic Behavior and Organization*, 182, pp. 472-492. doi: 10.1016/j.jebo.2019.03.006

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/21841/>

Link to published version: <https://doi.org/10.1016/j.jebo.2019.03.006>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

Do Actively Managed US Mutual Funds Produce Positive Alpha?

Rong Huang*

Keith Pilbeam[†]

William Pouliot ^{‡§}

Abstract

Using more general forms of equilibrium asset pricing models, we re-examine the recent literature on actively managed US Mutual Fund performance over the period 1984–2015. Using the false discovery technique, we show that the existing literature which is based upon unconditional versions of these models have underestimated performance of actively managed US funds because they produce inconsistent estimates of fund alpha coefficients. Our estimations of abnormal returns using conditional models allow the parameters that underlie the equilibrium asset pricing models to change which produces consistent estimates of funds alphas. We find that when returns are measured net of management and trading costs between 2.9% to 8.4% of US actively managed funds provide positive-alpha. This contrasts with existing studies that find no significant percentage of US Mutual Funds produce positive-alpha. We also find that different investment styles have significantly different percentages of positive-alpha funds.

Keywords: Mutual Funds, Capital Asset Pricing Model; CUSUM test, Bootstrap.

Classification: C3, C4, G1.

1 Introduction

The US mutual fund industry had, as of year-end 2015, a total value of nearly \$16 trillion assets under management much of which is actively managed. The industry itself has grown significantly, in 1984 there were just under 1,243 mutual funds while in 2015 there were over 8,115 with net assets rising from \$370 billion in 1984 to \$15,650 billion in 2015. Household ownership of mutual funds has risen from 10.8 million in 1984 (14.7% of households) to 54

*Business School University of Nottingham, Nottingham NG8 1BB United Kingdom, email: Rong.Huang@nottingham.ac.uk, ph: +44 (0)1158232546.

[†]Department of Economics City, University of London, London EC1V 0HB United Kingdom, email: K.S.Pilbeam@city.ac.uk, ph: +44 (0)2070400258.

[‡]Department of Economics, University of Birmingham, Edgbaston Birmingham B15 2TT United Kingdom, email: W.Pouliot@bham.ac.uk, ph: +44 (0)121 4145023.

[§]The authors would like to thank two anonymous referees, Ansgar Belke, Georgios Chortareas, Keith Cuthbertson, Dimitrios Asteriou, Josha Beckmann, Luca Riccetti, Jose Olmo, Tommaso Gabrieli, attendees at the ISCEF 2018 conference in Paris, those at the European Economic and Finance Society 2017 conference in Ljubljana and seminar participants at City, University of London for their valuable comments.

million in 2015 (43.7% of households)¹. Until recently, roughly half the total assets under management were held by actively managed mutual funds. They promote themselves as generating higher returns through their active trading compared to passively managed funds. Companies, such as Morningstar, Forbes, Dow Jones & Company and Zacks, offer rankings of mutual funds based on their returns and their particular fund's investment style. Such rankings are closely watched by investing households and investment flows tend to follow the rankings of the most successful mutual funds. The sheer size, value and importance of this industry to the US economy merits attention. In addition, in recent years there has been a significant shift towards more passive fund management in the form of index tracking funds and Exchange Traded Funds motivated by their significantly cheaper management expenses and trading costs as well as evidence that active fund managers have been underperforming passive investment returns.

In a well-known study, Barras et al. (2010) (hereafter BSW) estimate that just 0.6% of active US Mutual Funds have a true positive alpha net of management expenses and trading costs², with their point estimate being statistically insignificant. Their use of unconditional models to estimate abnormal fund performance (the alpha of a unconditional equilibrium asset pricing models) can be problematic because the estimate of alpha will be smaller than its true value. As such, our first and most important contribution is to show how estimates of the percentage of positive-alpha funds depend crucially upon whether conditional or unconditional models are employed. A result that arises because unconditional models, as used by BSW and Fama and French (2010) as well as others, tend to underestimate the value of a fund's alpha. Our second contribution consists of developing a nonlinear asset pricing model that makes no unnecessary assumptions on the appropriate information variables to use or on the relationship between the market risk factor and these variables. Hence our estimates of each mutual fund's alpha is consistent regardless of which of the unconditional/conditional three or four factor capital asset pricing models (hereafter 3F-CAPM or 4F-CAPM respectively) that we analyze is best. The third contribution is to update existing studies to 2015. Our final contribution is to evaluate fund performance within style categories, that is, we are able to see how the percentage of positive-alpha coefficients varies according to the investment style of the Fund distinguishing between aggressive growth, growth and growth & income funds.

The semi-strong form of the efficient markets hypothesis has two implications that are of interest to us here. The first is to posit that fund's market risk factor as measured by its beta can change because fund managers adjust their portfolios in the light of publicly available information. This theoretical possibility is confirmed via our statistical tests which show that 67.8% of funds' market risk beta change due to use of publicly available information. This suggests that unconditional models as employed by BSW are not always appropriate to measure a fund's returns. To estimate model parameters consistently, leads us to estimate conditional versions of these models. Here, model parameter are assumed to be a linear function of information variables that ensures consistency of parameter estimates whether or not a fund's market beta is time varying or not. The second implication is that the use of publicly available information by fund managers should not be considered as part of a fund's alpha. As such,

¹See *The investment company fact book* (2015)

²In their study Barras et al. (2010) refer to skilled, unskilled and zero skill fund managers. We prefer to use the terms negative, zero and positive alpha funds as the managers of the mutual fund can change over time.

we need to deduct from the alpha of unconditional models for the use of this information. For this reason, we adopt the models of Ferson and Schadt (1996) and Christopherson et al. (1998) to model the returns in our Mutual Fund dataset. The latter model nests Ferson and Schadt (1996) as well as the Carhart (1997), and has the advantage of producing consistent estimates of the alphas and betas even though the model itself may not apply to every fund. This feature is particularly useful in our context, as it can be applied to all funds' returns regardless of whether the market risk factor is time varying or not which is well suited since our tests reveal, in 67.8% of regressions, that the information variables affect the alpha and beta parameters.

The Christopherson et al. (1998) model only allows for two of the five parameters of the model to be time varying. It also assumes a linear relationship between the two parameters and the publicly available information variables, this assumption is unlikely to hold for all the return series in our dataset. To overcome these two issues, we propose a non-linear model with time varying parameters that is based on structural breaks in all or some of the parameters. As such, our model allows for more parameter changes which will likely result in a higher estimate of the proportion of funds producing positive-alpha; this is confirmed in our empirical analysis.

The rest of this paper is arranged as follows: Section 2 provides a brief summary of some of the previous research on fund performance; Section 3 discusses four models used to measure abnormal fund performance including our proposed non-linear version; Section 4 examines how the false discovery rate as employed by Barras et al. (2010) can be used to estimate the number of funds achieving a positive or negative alpha; Section 5 discusses the difference between net and gross returns and how they change the interpretation of a fund's alpha. It also explains the mutual fund data used to estimate the equilibrium asset pricing models in our study; Section 6 estimates performance using unconditional and conditional models and then details estimates of the percentages of negative, zero and positive-alpha funds; Section 6 estimates performance using unconditional and conditional models and then details estimates of the percentages of negative, zero and positive-alpha funds; Section 7 estimates fund performance using unconditional and conditional models and then details the estimates of the percentage of negative, zero and positive-alpha funds when fund performance is estimated via a nonlinear version of Christopherson et al. (1998); Section 8 provides estimates of alphas based upon fund investment styles; Section 9 concludes. Finally, Appendix A performs some robustness checks³.

2 Research on mutual fund performance

The US has a well developed mutual fund industry which is the largest amongst countries based on assets under management. Within this industry, actively managed funds represent a significant proportion of this total, and they have higher expense ratios and trading costs due to additional costs associated with active management. Motivated by concern over the value of higher fund costs to investors, a number of articles, published primarily in academic journals, have, with varying degrees of success, tried to assess performance of actively managed mutual funds. One of the earliest studies to contribute to this literature was Jensen (1968). By

³Appendix A contains two tables, Tables 9 and 10, of estimates on percentages of negative-, zero- and positive-alpha funds using the 90-day T-bill rate as the risk free, and also using the 3F-CAPM respectively.

adapting the capital asset pricing model (CAPM), he estimates this model on net returns series of 115 actively managed funds that existed during 1945 to 1964. As long as certain assumptions hold, the intercept of this model should be positive to reflect a manager's forecasting ability. If a manager can accurately forecast share prices, they will purchase shares whose price will rise and sell those whose price will fall. This implies the idiosyncratic component of the CAPM will have a positive mean and therefore its intercept will be positive. Jensen therefore argues that the intercept of this model (the alpha) is a measure of excess returns brought about by superior forecasting skill of a fund manager. Later on, Dybvig and Ross (1985) provide formal justification for its use as a measure of performance.

Results reported by Jensen shows that of the 115 funds he used, 97 had estimated alphas that were insignificantly different from zero, 3 had significant positive alphas, while 16 had significant negative alphas. Because the average value for the estimated alpha was -0.011 , he concludes managers are not able to forecast future stock prices well enough to cover management, research and transaction fees. Henriksson (1984) finds that estimated alphas, before load fees but net of expenses and trading costs, are not significantly different from zero. A result that implies managers have access to private information which allows them to generate positive abnormal returns to offset their cost.

In his study, Ippolito (1989), using yearly mutual fund data, found that 127 of 143 funds had estimated alphas insignificantly different from zero, 4 had significant negative alphas, while 12 had significant positive alphas. Seven funds, at most, are expected to have significant positive or negative alphas by chance alone. The five remaining alphas are not lucky but generate truly abnormal returns. He also calculates the mean of these estimated alphas and finds it to be positive and statistically significant regardless of benchmark (S&P 500, NYSE and S&P 500 and Salomon bond index) used to estimate the asset pricing equation. All this leads him to conclude that funds with higher turnover, fees and expenses earn sufficient risk-adjusted returns to offset their higher costs.

Grinblatt and Titman (1989), also looked at performance of mutual using an eight-portfolio benchmark. Regarding average fund performance using gross returns, they estimate average abnormal fund return to be 0.09% which is statistically insignificant. When funds are separated by investment objective, they find positive abnormal performance for aggressive-growth funds ($3.2\% = 0.027 \times 12$ annual return cf. page 404) and for growth funds ($1.9\% = 0.16 \times 12$ annual return cf. page 404), with both estimates statistically significant. They also employ F tests to determine jointly whether all 527 regression intercepts are zero and also whether they are all equal. Both tests reject these hypotheses. According to this evidence, managers of aggressive growth funds and growth funds generate superior performance.

Research in the 1990s tested for the presence of skilled managers based on tests of persistence of fund performance. Grinblatt and Titman (1992) test for persistence in fund abnormal returns. Their 10 year sample of fund returns data is split into two five-year subperiods. For each subperiod, fund abnormal returns are estimated. Then the slope coefficient from a cross-section regression of fund abnormal returns is computed by regressing the last five years of abnormal return on estimates of abnormal performance from first five years of returns data. A similar procedure is also applied to 109 passively formed portfolios to test for persistence in randomly formed portfolios. The estimated slope from the former regression turned out to be positive and statistically significant but not for the latter regression which leads them to conclude past performance affects future performance, clear evidence that performance persists.

Concerned that their test may only capture effects of persistence in fees and transaction costs, they rank abnormal returns into the 10% best and worst performing funds during their earlier five-year period. They then regress abnormal performance of best and worst funds on their abnormal performance in the next year. For the top performing funds, they find a positive and statistically significant relationship but a negative and statistically significant relationship for the worst performing funds. Persistence cannot be due to fees and other costs because abnormal returns of the best performing funds would fall by these costs.

Subsequently, Elton et al. (1993) show that the positive conclusions made by Ippolito are due to differences in the returns of non-S&P 500 stocks held by mutual funds during 1965 to 1984, the period used by Ippolito, compared to their returns during 1945 to 1964, the period used by Jensen. Once they account for the performance of non-S&P 500 stocks, they find the average estimated alpha to be negative and statistically significant (see their Table 5 page 10).

Hendricks et al. (1993) test the semi-strong form of market efficiency through tests for predictability of residuals obtained from estimating the four factor asset pricing model. They find residuals lagged by four months are correlated with contemporaneous residuals and conclude there is positive persistence in performance from one quarter to the next. In tests of persistence due to hot hands, they find its impact reaches a maximum when portfolios are ranked based on four quarter evaluation periods.

More recently Kosowski et al. (2006) introduce a data intensive resampling procedure, the bootstrap, to their search for skilled US fund managers. It is used because the cross-section distribution of fund alphas is too complex to be well approximated by asymptotic theory. Furthermore, the bootstrap is used to control for luck which previous studies had been unable to do. Not all abnormal fund performance can be attributed to skill as some funds with small p -values have positive alpha simply by chance. Using the bootstrap, they conclude most managers cannot compensate for management expenses and trading costs but a small subgroup more than compensate for these costs. They also find superior past performance can be used to predict future performance only for growth funds. Like the research of this paper, they also estimate conditional models of fund performance to obtain estimates of fund alphas and conclude that their results are not sensitive to model specification. Nevertheless, unconditional and conditional models produced different percentiles for the lower tail of the cross section distribution of fund alphas. Cuthbertson et al. (2008) adopt the bootstrap procedure as well which they use to find 12 funds that exhibit levels of performance not attributed to luck (cf. page 614).

In contrast to the methodology of Kosowski et al. (2006), Berk and van Binsbergen (2015) argue that it is superior to use a measure of value added based on the size of a fund to measure managerial skill rather than alpha based on gross or net returns. The reason for this lies with efficient capital markets. Investors will allocate funds to managers with the most skill and such funds will get larger over time, especially if there is persistence in managerial skill. Hence, a fund manager that earns an alpha based on gross returns of 40 basis points on a fund of \$10 billion is worth more than a fund manager earning an alpha of 100 basis points based on gross returns while managing \$100 million. In such circumstances, they point out that alpha based on gross returns is only an adequate measure of managerial skill if all mutual funds are of the same size. Berk and van Binsbergen (2015) also show that the largest funds typically have the best dollar value added so that even though a majority of funds have a negative alpha based

on gross returns, the average fund nonetheless has a \$3.2 million value added. They find that since managers fees are related to the success in producing value added, that even when alpha based on gross returns is positive, the alpha based on net returns is not significantly different from zero. It appears that skilled managers capture most of the alpha based on gross returns by increased management fees leaving no positive alpha for investors. This should be borne in mind when looking at our reported results, although the main focus of our study is correctly calculating the percentage of funds producing positive alpha based on gross and net returns.

3 Models to measure mutual fund performance

This section discusses four versions of a four factor capital asset pricing model, three of which are widely used in the literature to measure fund performance, while the fourth is our proposed non-linear extension of all three. The models cover constant parameters, time-varying parameters and non-linear parameters. According to Theorem 5 of Dybvig and Ross (1985), the alpha of these models can be used to capture a fund's abnormal performance. Since our estimates of the percentages of negative, zero and positive-alpha U.S. actively managed mutual funds are based on estimates of the intercept of the regressions it is crucial that they be estimated consistently. The literature on evaluating fund performance uses different versions of equilibrium asset pricing models to measure performance.

In the earliest studies, a one factor version of this model was employed where the beta was assumed to be constant through time. In his original study, Jensen (1968) had to make assumptions on a fund manager's behaviour to ensure that it would be constant. In particular, he assumed fund managers maintain a given level of portfolio risk. In practice, a mutual fund's level of risk is not constant because an active manager can switch between less and more risky equities or alter their weight within a fund (see Jensen (1968) p.394). Without this assumption, alpha cannot accurately measure a fund's excess return. Ippolito (1989) goes one step further and tests Jensen's assumption on returns of 143 mutual funds over the period 1965-1984. To do this, he separates the data on fund returns into two non-overlapping 10-year periods and then tests the constant beta assumption. He concludes for "15 out of 143 funds the hypothesis of stable betas could not be accepted". At a 5% significance level, seven funds would wrongly be concluded to have unstable betas over the twenty years. For the remaining eight, the beta would have changed from one period to the next.

More recently, papers by Barras et al. (2010) and Fama and French (2010) use a four factor version of this model which we refer to as the 4F-CAPM. Like Jensen, they use models in which the parameters are assumed to be constant throughout a fund's return history. Below, we discuss each of our four models used in this study.

3.1 Model 1: Carhart with constant betas

This version of the CAPM model was developed by Carhart (1997) who used it to estimate alpha; his model is reproduced below:

$$R_{i,t} - R_t^f = \alpha_i + \beta_{1,i}(R_t^m - R_t^f) + \beta_{2,i}SMB_t + \beta_{3,i}HML_t + \beta_{4,i}MOM_t + \varepsilon_{i,t}. \quad (1)$$

Here $R_{i,t}$ is the month t return of fund i , R_t^f is the risk free rate (proxied by the 30 day T-bill beginning-of-month yield); R_t^m is the month t return on the CRSP (Center for Research in Security Prices) NYSE/AMEX/NASDAQ value-weighted market portfolio; and SMB_t , HML_t and MOM_t are the month t returns on zero-investment factor mimicking portfolios for size, book-to-market and momentum obtained from Kenneth French's website. The three factor version of this model (referred to as the 3F-CAPM) removes the MOM_t from the regression.

3.2 Model 2: Ferson and Schadt (1996) with time-varying betas

To rectify the problem of constant parameter models, time-varying parameters have been used to evaluate fund performance. One model satisfying this condition is the 4F-CAPM of Ferson and Schadt (1996), a model which is based on three assumptions: (i) the functional form of asset pricing models used to describe conditional expected returns of assets; (ii) the use of publicly available information; and (iii) the functional form for the portfolio beta. Regarding the first assumption, Ferson and Schadt (1996) use a multiple-factor asset pricing model. As to the second, it refers to the notion of market efficiency. Ferson and Schadt (1996) state that use of unconditional models to measure performance can lead to abnormal performance by fund managers who use publicly available information. To control for common variation in parameters as well as risk premia caused by the use of public information, the betas must be time-varying.

Ferson and Schadt (1996) model the combined effect of these factors on the betas as "reduced form". Justification for this assumption comes from Admati et al. (1986) who assumes fund managers maximize a constant absolute risk aversion expected utility function under which a fund's return, as well as public information variables, are jointly normally distributed. Under this specification, the portfolio weights are linear functions of the information variables. As long as the risks associated with the underlying equity are constant, their reduced form specification follows. If this specification is reasonably accurate and appropriate information variables are used, then it is well known that least squares estimates of the parameters are consistent. These estimates would remain consistent even when the 4F-CAPM of equation (1) correctly models returns. This is important to remember when constructing accurate estimates of percentages of truly negative, zero and positive-alpha funds, as we do here.

Imposing the three assumptions of Ferson and Schadt (1996) on the 4F-CAPM, they arrive at a linear form for their conditional factor model which is given below:

$$R_{i,t} - R_t^f = \alpha_i + \beta_{1,i}(R_t^m - R_t^f) + \beta_{2,i}SMB_t + \beta_{3,i}HML_t + \beta_{4,i}MOM_t + B'_{5,i}(Z_{t-1} - \bar{Z})(R_t^m - R_t^f) + \varepsilon_{i,t}. \quad (2)$$

Here Z_{t-1} denotes a $J \times 1$ vector of predictive variables measured at the end of month $t - 1$, \bar{Z} is their mean and $B_{5,i}$ is the $J \times 1$ vector of coefficients. In equation (2), the intercept is assumed to be constant which is consistent with the model of manager skill developed by Berk and Green (2004). Our choice of conditioning variables follows BSW who use the 1-month T-bill yield; the dividend yield for the value-weighted NYSE/Amex stock index; the term spread; and the default spread.

3.3 Model 3: Christopherson et al. (1998) with time-varying alpha and betas

Further work by Christopherson et al. (1998) change the assumption of constant manager performance to allow it to be time-varying. They argue that not only do the betas need adjusting for use of public information but the alpha as well. Following the reduced form specification of Ferson and Schadt (1996), they set $\alpha = \alpha' + B'_{6,i}(Z_{t-1} - \bar{Z})$. After substitution in (2), this yields equation (3).

$$R_{i,t} - R_t^f = \alpha'_i + \beta_{1,i}(R_t^m - R_t^f) + \beta_{2,i}SMB_t + \beta_{3,i}HML_t + \beta_{4,i}MOM_t + B'_{5,i}(Z_{t-1} - \bar{Z})(R_t^m - R_t^f) + B'_{6,i}(Z_{t-1} - \bar{Z}) + \varepsilon_{i,t} \quad (3)$$

3.4 Model 4: A Nonlinear Regression Specification

While the conditional model of Christopherson et al. (1998) provides formal justification for the presence of time-varying alpha as well as betas in the 4F-CAPM there is additional empirical evidence of parameter instability. As previously discussed, Ippolito (1989) provides clear evidence of changing betas in enhanced versions of the CAPM. This is not without consequence for the CAPM estimated here because the reduced form specification used by Ferson and Schadt (1996) requires the risks of assets underlying the fund to be constant. Berndt (1991), who summarizes a large body of work on time-constant betas in these models, concludes the market risk beta is constant over five year spans only. Beyond this time frame, they may change. For example, the deregulation in the 1980s of the airline and telecommunication industries led to significant changes in the market risk of equity associated with these industries. In their study, Grout and Zalewska (2006) demonstrate empirically that there is a clear link between regulatory changes and market risk for a set of UK privatized companies over the period 1993-2000.⁴ Hendricks et al. (1993) also look at nonlinear models because they can arise from the nature and characteristics of equities held by funds and the market-timing ability of managers.⁵ In their study, Mamaysky et al. (2008) show that while static models may be true for some funds but clearly not for all funds, this is because investors pay active fund managers to move funds between different sectors and companies as part of a dynamic optimising strategy. They argue that the use of a Kalman filter can help detect managers with positive-alpha because of the Kalman filter model's ability to adapt to a funds' current loading on market risk in a manner that simple OLS regressions cannot. In particular, they find that as much as 20% of active fund managers may have market-timing ability.

Correct specification of the equilibrium asset pricing model is important because, without it, estimates of the risk factors will be inconsistent and estimates of a fund's alpha will underestimate its performance. To obtain consistent estimates of all parameters in the 4F-CAPM, it should be considered nonlinear in all parameters. The simplest nonlinear model that one can consider could be constructed by testing for changes in the parameters of the 4F-CAPM of Carhart (1997) using well-known tests for structural breaks. The model we outline here

⁴In particular they find that, other things being equal, a shift to profit sharing should reduce the risk of the regulated company to a fraction of the risk in the absence of the shift to profit sharing.

⁵Hendricks et al. (1993) page 112.

exploits nonparametric tests developed by Pouliot (2016)⁶. In the context of testing for a change only in the alpha of equation (1), the measure of fund performance, the CUSUM test is based on the residuals, $\hat{\varepsilon}_1, \dots, \hat{\varepsilon}_T$ from estimating this model. The statistic for a change in the alpha is given by equation (4):

$$\max_{1 < k < T} \left| \frac{\sum_{t=1}^k \hat{\varepsilon}_t}{\sqrt{k}} - \frac{\sum_{t=k+1}^T \hat{\varepsilon}_t}{\sqrt{T-k}} \right|; \quad (4)$$

and the test for change in betas of equation (1) is based on the statistic given by equation (5):

$$\max_{1 < k < T} \left| \frac{\sum_{t=1}^k \hat{\varepsilon}_t^2}{\sqrt{k}} - \frac{\sum_{t=k+1}^T \hat{\varepsilon}_t^2}{\sqrt{T-k}} \right|. \quad (5)$$

The statistic in equation (4) must be made robust to a possible change in the betas of equation (1), while the statistic in equation (5) must be made robust to a change in the alpha of equation (1).⁷ The date of change, t^* , must be estimated as well.⁸

These CUSUM tests can also be implemented to detect multiple changes. To do this, we follow the hierarchic binary segmentation algorithm of Vostrikova (1981), which begins with a test statistic, like the CUSUM in this context, to detect one change in the parameters over the entire sample. If the null of no change cannot be rejected, we conclude no change has occurred. If the tests reject the null hypothesis of no change in parameters, then the date of the change is estimated. The sample is divided into two segments: one consisting of data up to the estimated date of change, and the second consisting of data after the estimated date of change. Then the tests are applied to these two segments to determine whether an additional change occurs in each segment. If no additional change is detected, then the algorithm stops. However, if the test rejects the null, we estimate the date of change within the segment, the data is separated further and tests are applied again. The test is implemented sequentially in this way until no further changes are found.

We now detail how the CUSUM test can be used, along with an estimate of the date of change, to construct a nonlinear model⁹. First, the CUSUM test for one change in the parameters is performed. If it does not reject the null hypothesis of no change, then equation (1) is estimated. If, however, the null hypothesis of no change is rejected, then based on the estimated date of the change, the sample is divided into two segments: one segment up to the estimated date of the change, and a second segment after this date. Hence, there would be two equations corresponding to equation (1) and the following regression model is estimated:

$$R_{i,t} - R_t^f = \begin{cases} \alpha_i^{(1)} + \beta_{1i}(R_t^m - R_t^f) + \beta_{2i}SMB_t + \beta_{3i}HML_t + \beta_{4i}MOM_t + \varepsilon_{i,t} & t_0 \leq t \leq \hat{t}^* \\ \alpha_i^{(2)} + \beta_{1i}(R_t^m - R_t^f) + \beta_{2i}SMB_t + \beta_{3i}HML_t + \beta_{4i}MOM_t + \varepsilon_{i,t} & \hat{t}^* + 1 \leq t \leq t_T. \end{cases} \quad (6)$$

⁶Since that paper discusses the tests and modelling framework in detail, so we provide only a brief review.

⁷A more formal explanation of how this is achieved is provided in sections 4.1 and 4.2 of Pouliot (2016).

⁸Equation (16) of Pouliot (2016) provides one such estimator.

⁹Throughout the CUSUM tests, the significance level is set to 5%.

Instead, if there is only one change in the slope and no change in intercept¹⁰, the following model is estimated:

$$R_{i,t} - R_t^f = \begin{cases} \alpha_i + \beta_{1i}^{(1)}(R_t^m - R_t^f) + \beta_{2i}^{(1)}SMB_t + \beta_{3i}^{(1)}HML_t + \beta_{4i}^{(1)}MOM_t + \varepsilon_{i,t} & t_0 \leq t \leq \hat{t}^* \\ \alpha_i + \beta_{1i}^{(2)}(R_t^m - R_t^f) + \beta_{2i}^{(2)}SMB_t + \beta_{3i}^{(2)}HML_t + \beta_{4i}^{(2)}MOM_t + \varepsilon_{i,t} & \hat{t}^* + 1 \leq t \leq t_T. \end{cases} \quad (7)$$

For more complicated cases, if there is one change in the intercept and one change in the slope, the following model is estimated:

$$R_{i,t} - R_t^f = \begin{cases} \alpha_i^{(1)} + \beta_{1i}^{(1)}(R_t^m - R_t^f) + \beta_{2i}^{(1)}SMB_t + \beta_{3i}^{(1)}HML_t + \beta_{4i}^{(1)}MOM_t + \varepsilon_{i,t} & t_0 \leq t \leq \hat{t}_1^* \\ \alpha_i^{(2)} + \beta_{1i}^{(1)}(R_t^m - R_t^f) + \beta_{2i}^{(1)}SMB_t + \beta_{3i}^{(1)}HML_t + \beta_{4i}^{(1)}MOM_t + \varepsilon_{i,t} & \hat{t}_1^* + 1 \leq t \leq \hat{t}_2^* \\ \alpha_i^{(2)} + \beta_{1i}^{(2)}(R_t^m - R_t^f) + \beta_{2i}^{(2)}SMB_t + \beta_{3i}^{(2)}HML_t + \beta_{4i}^{(2)}MOM_t + \varepsilon_{i,t} & \hat{t}_2^* + 1 \leq t \leq t_T \end{cases} \quad (8)$$

A similar process is followed when we have more than two parameter changes. For brevity, we only list the above three cases. As can be seen, after a change point is detected, fund return history is divided into subsamples according to the estimated date of the change and one regression is estimated on each of these subsamples¹¹. It should be noted that there might be changes very close to a previous change or located near the beginning or the end of the sample, leaving only a few observations in the subsamples. If the subsample contains too few observations to obtain accurate estimates, we assume there is no change.

Our proposed CUSUM tests have four main advantages in the context of our study. First, they are able to detect more than one parameter change within a linear regression model using binary segmentation. This is useful given that some funds may have multiple changes in their factor loadings and alpha. Second, the test allows us to distinguish a change in the slope parameters from a change in the intercept while remaining robust to each other. This is particularly important as the intercept and the slope have different interpretations in our context, in that it tells us which changes are due to time-varying alpha and which ones occurred because of shifts in the factor loadings. Third, the test has power against parameter

¹⁰It should be noted that although the test is able to distinguish a change in the alpha from one in the betas, it cannot tell which one of the regression betas has changed. Hence, when there is a change in the slope, we assume all the regression slopes change. This setting still gives us consistent estimates of the betas even when only a subset of the betas change but now the estimates are not efficient.

¹¹The CUSUM tests implemented here do not control the experiment-wise error rate at the significance level of 5% which is the significance level used for these tests. We expect this to have little impact on our estimates of model parameters. Falsely concluding that parameters of the 4F-CAPM have changed means that the mutual fund returns series is split into two parts and two models are then estimated: one before the change and one after. Since the CUSUM tests have falsely rejected the null-hypothesis of no change, estimates produced from estimating two models remain unbiased and consistent for the parameters of the 4F-CAPM but are no longer efficient. The only consequence from our choice of one significance level is that the experiment-wise error rate will be large than 5% and therefore Table 4 will report more changes than the true number of changes that took place. Therefore the number of changes reported there represent an upper bound on the true number of changes.

changes that are located near the limits of the sample period, which implies that the tests can capture changes that occur early or late in the alpha and betas. Fourth, no assumption is made on how information variables may affect the betas and the alpha, or on the correct set of information variables to use. Hence, the nonlinear model produces estimates of all parameters that are consistent regardless of these two specifications. Moreover, should the 4F-CAPM of equation(1) be correct, estimates of the parameters of the nonlinear model remain consistent.

4 Using the false discovery rate to estimate the proportions of negative, zero and positive alpha funds

We follow the statistical methodology proposed by Barras et al. (2010) to calculate the percentages of positive alpha producing funds. Correct evaluation of abnormal fund performance requires control for the possibility that a fund may be lucky or unlucky. There are funds that have no real alpha but are incorrectly classified as having either a positive or negative alpha. Incorrect classification arises because a fund's abnormal performance. These false rejections are called false discoveries which must be taken into account to obtain an accurate estimates of the percentages of positive and negative-alpha funds. To do this Barras et al. (2010) use the false discovery rate procedure.

Fund performance is based on the alpha of each fund in the population. A fund that performs well on a risk-return basis is a fund where the alpha is positive and poorly performing if it has a negative alpha, otherwise the fund has zero alpha. Since alpha is a parameter, it cannot be observed. An estimate is therefore needed which is obtained by estimating the 4F-CAPM using data on fund returns and the four factors. Using the estimated funds alphas we can construct a t-statistic and then choose a significance level and observe whether a fund's estimated alpha lies outside the critical value implied by that significance level. To assess the performance of a single fund, the null hypothesis is that the fund has a zero-alpha. versus the alternative that the fund provides either a positive or a negative alpha:

$$H_0 : \alpha_i = 0 \text{ versus } H_a : \alpha_i \neq 0; \quad (9)$$

In the above case, the Type I error (false positive) is controlled at a predetermined significance level. To assess performance of N funds, this test should be simultaneously applied across all of them. This leads to a multiple hypothesis testing situation which is given in (10):

$$\begin{aligned} H_{0,1}: \alpha_1 = 0 \text{ versus } H_{a,1}: \alpha_1 \neq 0, \\ \vdots \\ H_{0,N}: \alpha_N = 0 \text{ versus } H_{a,N}: \alpha_N \neq 0. \end{aligned} \quad (10)$$

To account for the false discoveries in their calculations, the process begins with the choice of a significance level γ . Then the probability of a true zero-alpha fund exhibiting luck equals $\gamma/2$. Given the symmetry of each test, it also implies that the probability of a true zero-alpha fund being unlucky is $\gamma/2$. The proportion of true zero-alpha funds in the population of N funds is π_0 , an unknown parameter. This parameter is used to calculate the expected proportion of true zero-alpha funds being lucky ($E(F_\gamma^+)$) or unlucky ($E(F_\gamma^-)$) which is given by equation

(11):

$$E(F_\gamma^+) = E(F_\gamma^-) = \pi_0 \cdot \gamma/2. \quad (11)$$

Knowledge of the expected proportion of lucky funds allows one to calculate the proportion of truly positive-alpha funds by subtracting the proportion of lucky funds from the expected proportion of significant positive alpha funds ($E(S_\gamma^+)$). A significant positive alpha fund is determined by the p -value of its test statistic: the p -value of a one sided test must be less than or equal $\gamma/2$ and have a positive t -statistic. A similar definition applies to a significant negative alpha fund, except here the t -statistic is negative while the condition on the p -value remains the same. The expected proportion of positive-alpha funds, denoted by $E(S_\gamma^+)$, and negative-alpha funds, denoted by $E(S_\gamma^-)$, must be adjusted according to equations (12) and (13), respectively, to obtain the expected proportions of truly positive-alpha funds, $E(T_\gamma^+)$, and truly negative-alpha funds, $E(T_\gamma^-)$:

$$E(T_\gamma^+) = E(S_\gamma^+) - E(F_\gamma^+) = E(S_\gamma^+) - \pi_0 \times \gamma/2, \quad (12)$$

$$E(T_\gamma^-) = E(S_\gamma^-) - E(F_\gamma^-) = E(S_\gamma^-) - \pi_0 \times \gamma/2; \quad (13)$$

Using different levels of γ allows us to see how dispersed or concentrated positive-alpha funds are in the extreme right tail of the cross-sectional distribution of alpha. For instance, if the expected proportion of positive-alpha funds ($E(T_\gamma^+)$) increases slightly as we change the value of γ from 0.1 to 0.2, then most of the positive-alpha funds should lie in the extreme right tail since the majority of the additional significant funds are merely lucky. However, if positive-alpha funds are dispersed randomly throughout the right tail, then we would expect to see a more significant increase in $E(T_\gamma^+)$ as γ increases.

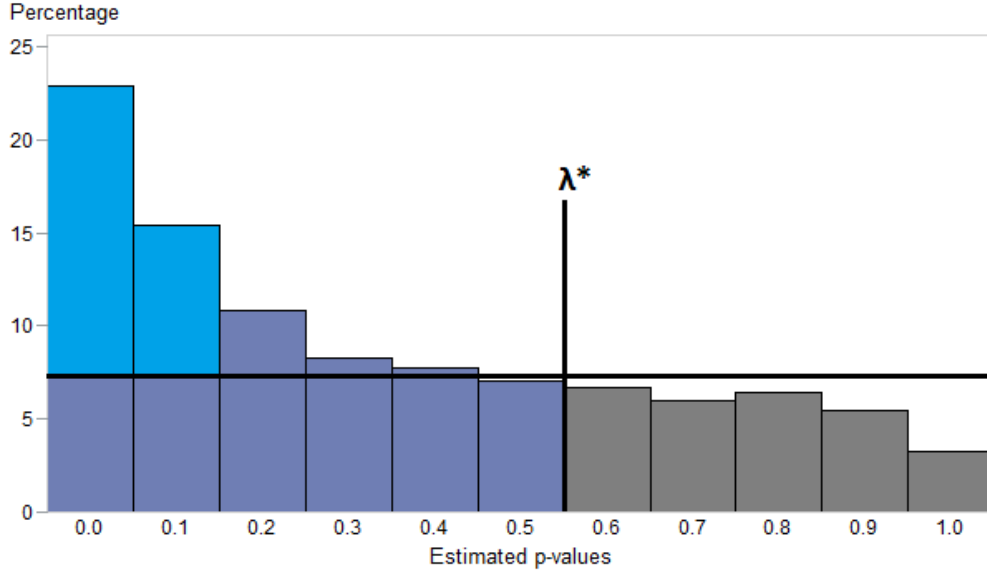
To understand some of the information and estimates that are reported in up coming tables, we provide the definition of the false discovery rate (hereafter FDR) used widely in this literature. Storey et al. (2004) defines the FDR as is given by equation (14)

$$FDR = E\left(\frac{F_\gamma}{S_\gamma}\right), \quad (14)$$

where F_γ and S_γ represent the proportions of falsely rejected null-hypotheses and rejected null-hypotheses out of all tests. According to this definition, we are interested in the following hypothesis $H_0 : \alpha_i = 0$ versus the two-sided alternative hypothesis $H_1 : \alpha_i \neq 0$ for $i = 1, \dots, N$. When testing for the presence of positive-alpha funds, the alternative becomes $H_1 : \alpha_i > 0$. In this situation, we define FDR^+ to be $FDR^+ = E\left(\frac{F_\gamma^+}{S_\gamma^+}\right)$ where F_γ^+ and S_γ^+ are the proportions of falsely rejected and rejected null hypotheses from all tests respectively. When testing for the presence of negative-alpha funds the alternative hypothesis becomes $H_1 : \alpha_i < 0$ and FDR^- is defined in a similar fashion.

4.1 Estimation of the proportion of zero-alpha funds

To estimate the above population proportions, the key is to estimate the proportion of zero-alpha funds (π_0) in the population of N funds. To estimate π_0 , we use the fact that truly negative-alpha or truly positive-alpha funds should have small p -values since their estimated

Figure 1: Histogram of Fund p -values

t -statistics tend to be very large which corresponds to the blue colour in Figure 1. Therefore, π_0 is estimated as follows:

$$\hat{\pi}_0(\lambda^*) = \frac{\hat{W}(\lambda^*)}{N} \times \frac{1}{(1 - \lambda^*)}, \quad (15)$$

where λ^* is a threshold (how to determine λ^* will be discussed later), after which the histogram of p -values as shown in Figure 1 becomes flat reflecting the fact that the p -values greater than the threshold (in grey) should all come from the group of truly zero-alpha funds whose p -value is uniformly distributed. Let $\hat{W}(\lambda^*)$ represent the number of funds whose p -value is greater than λ^* . Then $\hat{W}(\lambda^*)/N$ is the area of the histogram after the threshold λ^* . To extrapolate the area $(\hat{W}(\lambda^*)/N)$ over the entire interval from zero to one, we divide by $(1 - \lambda^*)$. To estimate π_0 , we have to compute the p -value associated with the null hypothesis in equation (9) for each fund. In fact, p -values of funds are the only inputs to estimate π_0 . Given that the finite sample distribution of the t -statistic is more likely to be asymmetric, a bootstrap procedure is used to compute fund p -values. Specifically, we compute fund p -values following the approach proposed by Davidson and MacKinnon (2004):

$$\hat{p}_i = 2 \times \min \left(\frac{1}{B} \sum_{b=1}^B I \left\{ \widehat{t}_i^{*b} > \widehat{t}_i \right\}, \frac{1}{B} \sum_{b=1}^B I \left\{ \widehat{t}_i^{*b} < \widehat{t}_i \right\} \right). \quad (16)$$

Where B is the number of bootstrap runs ($B=1000$), and $I \left\{ \widehat{t}_i^{*b} > \widehat{t}_i \right\}$ takes the value one if the bootstrapped t -statistic (\widehat{t}_i^{*b}) is greater than the t -statistic (\widehat{t}_i) .

Once the bootstrapped p -values are calculated, we implement the bootstrap approach proposed by Storey (2002) to choose the optimal λ and then to estimate the proportion of truly zero-alpha funds in the population using equation (15). To select the optimal λ , first, $\hat{\pi}_0(\lambda)$ is computed for each of the λ values ($\lambda=0.30, 0.35, 0.40, \dots, 0.70$). Then, we draw

with replacement from the $N \times 1$ vector of fund p -values to form 1000 bootstrapped samples. For each bootstrapped sample, the bootstrapped $\hat{\pi}_0^b(\lambda)$ ($b=1,2,\dots,1000$) is computed. Finally, the optimal λ (λ^*) is chosen to minimize the estimated Mean Squared Error (MSE), which is as follows:

$$\widehat{MSE}(\lambda) = \frac{1}{1000} \sum_{b=1}^{1000} \left[\hat{\pi}_0^b(\lambda) - \min_{\lambda} \hat{\pi}_0(\lambda) \right]^2. \quad (17)$$

Now that we have the estimated value of π_0 , all the statistics listed below can be computed:

$$\hat{F}_{\gamma}^{+} = \hat{F}_{\gamma}^{-} = \hat{\pi}_0 \times \gamma/2. \quad (18)$$

Then the estimated proportion of truly negative-alpha (\hat{T}_{γ}^{+}) and truly positive-alpha (\hat{T}_{γ}^{-}) funds with respect to the significance level γ is:

$$\hat{T}_{\gamma}^{+} = \hat{S}_{\gamma}^{+} - \hat{F}_{\gamma}^{+} = \hat{S}_{\gamma}^{+} - \hat{\pi}_0 \times \gamma/2, \quad (19)$$

$$\hat{T}_{\gamma}^{-} = \hat{S}_{\gamma}^{-} - \hat{F}_{\gamma}^{-} = \hat{S}_{\gamma}^{-} - \hat{\pi}_0 \times \gamma/2. \quad (20)$$

Finally, the estimated proportion of truly positive-alpha and truly negative-alpha funds in the population of N funds is given by:

$$\hat{\pi}_A^{+} = \hat{T}_{\gamma^*}^{+} \quad (21)$$

$$\hat{\pi}_A^{-} = \hat{T}_{\gamma^*}^{-}, \quad (22)$$

where γ^* is the optimal significance level chosen to have minimum MSE for $\hat{\pi}_A^{+}$ and $\hat{\pi}_A^{-}$. The optimal value is selected using a bootstrap approach which is similar to the one used to choose the optimal value of λ . For a range of γ values ($\gamma = 0.05, 0.10, \dots, 0.50$), we calculate the proportion of positive-alpha funds ($\hat{\pi}_A^{+}(\gamma)$). Then, we resample with replacement from the $N \times 1$ vector of p -values to produce 1000 bootstrap replications of $\hat{\pi}_A^{+}(\gamma)$. Finally, the estimated MSE of $\hat{\pi}_A^{+}(\gamma)$ is computed as:

$$\widehat{MSE}^{+}(\gamma) = \frac{1}{1000} \sum_{b=1}^{1000} \left[\hat{\pi}_A^{b+}(\gamma) - \max_{\gamma} \hat{\pi}_A^{+}(\gamma) \right]^2; \quad (23)$$

We choose optimal γ^{+} to minimize $\widehat{MSE}^{+}(\gamma)$. Accordingly, we also compute $\widehat{MSE}^{-}(\gamma)$ for each significance level γ . If $\min_{\gamma} \widehat{MSE}^{+}(\gamma) < \min_{\gamma} \widehat{MSE}^{-}(\gamma)$, then we choose γ to minimize $\widehat{MSE}^{+}(\gamma)$ and set $\hat{\pi}_A^{-}(\gamma^*) = 1 - \hat{\pi}_0 - \hat{\pi}_A^{+}(\gamma^*)$ to preserve the equality $1 = \hat{\pi}_0 + \hat{\pi}_A^{+} + \hat{\pi}_A^{-}$. Otherwise, we choose the optimal γ to minimize $\widehat{MSE}^{-}(\gamma)$ and set $\hat{\pi}_A^{+}(\gamma^*) = 1 - \hat{\pi}_0 - \hat{\pi}_A^{-}(\gamma^*)$. Estimation of standard errors of the above mentioned estimates are calculated based on the formula provided on page 2 of the Internet Appendix to Barras et al. (2010).

5 Measurement of positive-alpha fund returns on a net and gross basis and the mutual fund data

Following BSW, the mutual fund data used here is separated into different categories according to whether fund returns (gross or net), and also along fund investment styles. When funds returns are measured on a net basis, a positive alpha tells us that a fund has generated excess returns sufficient to cover all management expenses and trading costs, whereas a positive alpha based on gross returns indicates a fund's excess returns are not necessarily sufficient to cover management expenses. A fund can then be classified into one of three categories depending on a fund's performance as follows:

1. A fund with negative alpha on a net basis: does not cover trading costs and management expenses ($\alpha_{net} < 0$);
2. A fund with a zero-alpha on a net basis: just covers trading costs and management expenses ($\alpha_{net} = 0$);
3. Fund with positive alpha on a net basis: more than covers trading costs and management fees ($\alpha_{net} > 0$).

On the other hand, when only trading costs, but not management expenses, are used they are referred to as gross returns. these categories become defined as:

1. A fund with negative alpha on a gross basis: does not cover trading costs ($\alpha_{gross} < 0$);
2. A fund with a zero-alpha on a gross basis: just covers trading costs ($\alpha_{gross} = 0$);
3. A fund with a positive-alpha on a gross basis: more covers trading costs ($\alpha_{gross} > 0$).

The monthly U.S. open-ended domestic equity mutual fund return data used in this study comes from the Center for Research in Security Prices (CRSP) mutual fund survivor-bias free database. It covers the period from January 1984 to March 2015 and includes funds that existed at any time during this period. To focus on the performance of actively managed mutual funds, we follow BSW and include mutual funds as long as they are actively managed and invest primarily in U.S. common stocks. Passively managed funds like bond and money market as well as sector specific funds are therefore excluded from our study. To construct monthly fund returns, data was collected on beginning-of-month month fund returns by share class, assets under management (AUM), expense ratio and fund type. Monthly fund return was calculated by multiplying the return in each share class by its value weight and then adding these value-weighted returns to form net or gross fund returns. Gross fund returns are obtained by adding to net returns the expense ratio divided by 12. As in BSW, we require funds to have at least 60 months of return history to be included in our analysis.

Our data is survivor-bias free like that of BSW; however we do not go back to 1975 as in their study. During the period 1962 to 1983, 15% of funds on CRSP report annual returns rather than monthly returns. The average annual equal weight return for these funds is 5.29% lower than for funds that report monthly returns. Due to this issue, we follow Fama and French (2010) and use data that begins in 1984 rather than follow BSW who begin in January 1975. As in BSW, we investigate the performance of mutual funds with different investment

styles. For this purpose we split our data into three categories according to their self-declared investment styles¹².

1. Aggressive growth funds - are funds that aim to achieve the highest capital gains and generally have the greatest risk; they usually hold investments with high growth potential as well as great price volatility.
2. Growth funds - are funds that aim to achieve capital appreciation and often have above average risks; by investing in growth stocks instead of those paying dividends, they gain the benefit of large capital increases.
3. Growth & income funds - are funds that value both capital growth and dividend income; they invest in a range of products that are able to provide investors income through dividends as well as capital gains.

There are 3,673 funds in our sample of which 306 of our funds are classified as aggressive growth funds, 2458 are classified as growth funds and a further 1419 funds are classified as growth & income funds¹³.

6 Measuring fund performance using traditional and time-varying 4F-CAPM

This section estimates the percentages of negative, zero and positive-alpha funds using the false discovery rate (hereafter FDR). These estimates are recorded in the Tables 1, 2 and 3 that follow. Panel A of each table lists estimates based on net returns, while Panel B lists them based on gross returns. Results obtained in each of these tables are based on one of three models: Carhart (1997) is used for Table 1, Ferson and Schadt (1996) is used for Table 2 and Christopherson et al. (1998) is used for Table 3.

We now discuss the main results presented in these tables. Panel A of Table 1 shows the percentages of funds having negative, zero and positive-alpha on a net return basis, and reveals that only 0.8% have positive-alpha, 32% have negative-alpha and 67.2% of funds have zero-alpha. These rather damning results are similar to those of BSW who report 0.6% of funds have positive-alpha, 75.4% have zero-alpha and 24% have negative-alpha. Our estimates of the percentages of positive alpha funds of 0.8%, while similar to the 0.6% report by BSW, differs in that our estimate is statistically significant while their estimate is not.

When these percentages are estimated using gross returns in Panel B, the percentages of positive alpha funds is now much larger than the corresponding estimate based on net returns: 12.6% of funds have a positive-alpha, 80.3% of funds have zero-alpha and 7.1% of funds

¹²The CRSP dataset permits us to identify the investment objective of a fund for a certain year. If, however, we find a change in the style for a fund in a certain year, we split the fund's return history into two separate records according to the date their objective changed. Hence, we may have more funds when we divide the universe of sample into different investment categories.

¹³The number of funds from all three categories in total more than the number of funds in our full sample. This is because funds can change their investment style/objective during their fund history. When the whole universe of funds is split into different investment categories there are some funds that may have parts of their return history in two or even three different investment categories.

Table 1:

Percentages of Negative-, Zero- and Positive-Alpha Funds: 1984-2015 (Constant Parameter 4F-CAPM; Net and Gross Returns)

Performance of funds is estimated using the 4F-CAPM using equation (1) of Carhart (1997). Panel A (net returns) and Panel B (gross returns) show the estimated proportions of zero-alpha ($\hat{\pi}_0$), negative alpha ($\hat{\pi}_A^-$) and positive alpha ($\hat{\pi}_A^+$) funds using the FDR approach based on a dataset of 3,673 funds for net returns and gross returns. Each panel also shows significant funds ($\hat{S}_\gamma^-, \hat{S}_\gamma^+$) across the distribution of funds in the left tail and right tail using four different significance levels ($\gamma = 0.05, 0.1, 0.15, 0.2$). Furthermore, significant funds are decomposed into unlucky (\hat{F}_γ^-) and negative alpha (\hat{T}_γ^-) funds in the left tail, and lucky (\hat{F}_γ^+) and positive alpha (\hat{T}_γ^+) funds in the right tail. At the bottom of each panel, we show the FDR in the left and right tail at different significance levels. The sample covers the period from January 1984 to March 2015. Standard errors are indicated by parentheses and located below the estimate¹⁴.

Panel A: net returns, constant parameter model										
	Zero alpha ($\hat{\pi}_0$)				Negative alpha ($\hat{\pi}_A^-$)				Positive alpha ($\hat{\pi}_A^+$)	
Percentage	67.2				32.0				0.8	
	(0.02)				(0.02)				(0.00)	
Number of funds	2469				1176				28	
Impact of Luck in the Left and Right Tails										
	Left Tail				Right Tail					
Signif.Level(γ)	0.05	0.10	0.15	0.20	0.20	0.15	0.10	0.05	Signif.Level(γ)	
Signif. \hat{S}_γ^- (%)	14.0	19.5	24.6	28.5	7.3	5.9	4.3	2.5	(%)Signif. \hat{S}_γ^+	
	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)		
Unlucky \hat{F}_γ^- (%)	1.7	3.4	5.0	6.7	6.7	5.0	3.4	1.7	(%)Lucky \hat{F}_γ^+	
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)		
Negative alpha \hat{T}_γ^- (%)	12.3	16.1	19.6	21.8	0.6	0.9	1.0	0.8	(%)Positive alpha \hat{T}_γ^+	
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)		
FDR ⁻	12.0	17.2	20.5	23.6	91.8	85.3	77.6	68.6	FDR ⁺	
Panel B: gross returns, constant parameter model										
	Zero alpha ($\hat{\pi}_0$)				Negative alpha ($\hat{\pi}_A^-$)				Positive alpha ($\hat{\pi}_A^+$)	
Percentage	80.3				7.1				12.6	
	(0.01)				(0.01)				(0.01)	
Number of funds	2951				261				461	
Impact of Luck in the Left and Right Tails										
	Left Tail				Right Tail					
Signif.Level(γ)	0.05	0.10	0.15	0.20	0.20	0.15	0.10	0.05	Signif.Level(γ)	
Signif. \hat{S}_γ^- (%)	4.1	6.6	9.3	11.1	19.6	16.1	12.9	8.3	(%)Signif. \hat{S}_γ^+	
	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)		
Unlucky \hat{F}_γ^- (%)	2.0	4.0	6.0	8.0	8.0	6.0	4.0	2.0	(%)Lucky \hat{F}_γ^+	
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)		
Negative alpha \hat{T}_γ^- (%)	2.1	2.5	3.2	3.1	11.6	10.1	8.9	6.3	(%)Positive alpha \hat{T}_γ^+	
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)		
FDR ⁻	48.9	61.2	65.1	72.1	40.9	37.3	31.1	24.1	FDR ⁺	

have negative alpha. All of these estimates are statistically significant. BSW estimate of the percentage of positive-alpha funds is 9.6% which is smaller than ours by 3 percentage points, while 85.9% have zero-alpha which is larger by 5.2% than our estimate. They estimate the percentage of negative-alpha funds at 4.5% which is smaller than our 7.1% estimate.

In Table 2, using the conditional model of Ferson and Schadt (1996), we estimate the percentage of positive-alpha funds on a net basis to be 3.2% (25.1% for gross returns) and find it to statistically significant. This is larger than the 0.8% (12.6% for gross returns) reported in Table 1 and larger than the 0.6% percentage reported by BSW. The percentage

Table 2:

Percentages of Negative, Zero, and Positive-Alpha Funds: 1984-2015 (Conditional Model)

Performance of funds is estimated using the conditional 4F-CAPM of Ferson-Schadt. Panel A (net returns) and Panel B (gross returns) show the estimated proportions of zero alpha ($\hat{\pi}_0$), negative alpha ($\hat{\pi}_A^-$) and positive alpha ($\hat{\pi}_A^+$) funds using the FDR approach based on a dataset of 3,673 funds using net and gross returns. Each panel also shows significant funds (\hat{S}_γ^- , \hat{S}_γ^+) across the distribution of funds in the left and right tail using four different significance levels ($\gamma = 0.05, 0.1, 0.15, 0.2$). Furthermore, significant funds are decomposed into unlucky (\hat{F}_γ^-) and negative alpha (\hat{T}_γ^-) funds in the left tail, and lucky (\hat{F}_γ^+) and positive alpha (\hat{T}_γ^+) funds in the right tail. At the bottom of each panel, we show the FDR in the left and right tail at different significance levels. The sample covers the period from January 1984 to March 2015. Standard errors are indicated by parentheses and located below the estimate.

Panel A: net returns										
	Zero alpha ($\hat{\pi}_0$)				Negative alpha ($\hat{\pi}_A^-$)				Positive alpha ($\hat{\pi}_A^+$)	
Percentage	66.9				29.9				3.2	
	(0.02)				(0.02)				(0.00)	
Number of funds	2458				1097				118	
Impact of Luck in the Left and Right Tails										
	Left Tail				Right Tail					
Signif.Level(γ)	0.05	0.10	0.15	0.20	0.20	0.15	0.10	0.05	Signif.Level(γ)	
Signif. \hat{S}_γ^- (%)	13.9	19.3	24.1	27.9	9.5	7.8	5.8	3.8	(%)Signif. \hat{S}_γ^+	
	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)		
Unlucky \hat{F}_γ^- (%)	1.7	3.3	5.0	6.7	6.7	5.0	3.3	1.7	(%)Lucky \hat{F}_γ^+	
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)		
Negative alpha \hat{T}_γ^- (%)	12.2	15.9	19.0	21.2	2.8	2.8	2.5	2.2	(%)Positive alpha \hat{T}_γ^+	
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)		
FDR ⁻	12.0	17.4	20.8	24.0	70.2	64.0	57.4	43.6	FDR ⁺	

Panel B: gross returns										
	Zero alpha ($\hat{\pi}_0$)				Negative alpha ($\hat{\pi}_A^-$)				Positive alpha ($\hat{\pi}_A^+$)	
Percentage	69.8				5.1				25.1	
	(0.02)				(0.01)				(0.01)	
Number of funds	2565				187				921	
Impact of Luck in the Left and Right Tails										
	Left Tail				Right Tail					
Signif.Level(γ)	0.05	0.10	0.15	0.20	0.20	0.15	0.10	0.05	Signif.Level(γ)	
Signif. \hat{S}_γ^- (%)	4.3	6.6	8.8	11.0	22.7	19.5	15.9	11.1	(%)Signif. \hat{S}_γ^+	
	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)		
Unlucky \hat{F}_γ^- (%)	1.7	3.5	5.2	7.0	7.0	5.2	3.5	1.7	(%)Lucky \hat{F}_γ^+	
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)		
Negative alpha \hat{T}_γ^- (%)	2.6	3.1	3.6	4.0	15.7	14.3	12.4	9.4	(%)Positive alpha \hat{T}_γ^+	
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)		
FDR ⁻	40.3	53.0	59.2	63.5	30.8	26.9	21.9	15.7	FDR ⁺	

of negative-alpha funds is reported to be 29.2% (5.1% for gross returns) which is statistically significant and smaller than 32% (7.1% for gross returns) estimated using equation (1) but is larger than the percentage reported by BSW.

In Table 3, we present the results using the conditional model of Christopherson et al. (1998). Our estimate of the percentage of positive-alpha funds on a net return basis is 2.9% (20.4% for gross returns) which is statistically significant. The percentage of zero-alpha funds is 66.5% (76.1% for gross returns) while the percentage of negative-alpha funds is 30.6% on a net basis (3.5% on a gross basis).

Table 3:

Percentage of Negative, Zero and Positive-Alpha Funds: 1984-2015 (Conditional Alpha-beta Model)

Performance of funds is estimated using the conditional alpha-beta model Christopherson et al. (1998) using net and gross returns. Panel A (net returns) and Panel B (gross returns) show the estimated proportions of zero alpha ($\hat{\pi}_0$), negative alpha ($\hat{\pi}_A^-$) and positive alpha ($\hat{\pi}_A^+$) funds using the FDR approach. Each panel also shows significant funds ($\hat{S}_\gamma^-, \hat{S}_\gamma^+$) across the distribution of funds in the left and right tail using four different significance levels ($\gamma = 0.05, 0.1, 0.15, 0.2$). Furthermore, significant funds are decomposed into unlucky (\hat{F}_γ^-) and negative alpha (\hat{T}_γ^-) funds in the left tail, and lucky (\hat{F}_γ^+) and positive alpha (\hat{T}_γ^+) funds in the right tail. At the bottom of each panel, we show the FDR in the left and right tail at different significance levels. The sample covers the period from January 1984 to March 2015. Standard errors are indicated by parentheses and located below the estimate.

Panel A: net returns										
	Zero alpha ($\hat{\pi}_0$)				Negative alpha ($\hat{\pi}_A^-$)				Positive alpha ($\hat{\pi}_A^+$)	
Percentage	66.5				30.6				2.9	
	(0.02)				(0.01)				(0.02)	
Number of funds	2443				1124				106	
Impact of Luck in the Left and Right Tails										
	Left Tail				Right Tail					
Signif.Level(γ)	0.05	0.10	0.15	0.20	0.20	0.15	0.10	0.05	Signif.Level(γ)	
Signif. \hat{S}_γ^- (%)	13.5	19.4	24.4	28.3	9.5	7.7	5.7	3.8	(%)Signif. \hat{S}_γ^+	
	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)		
Unlucky \hat{F}_γ^- (%)	1.7	3.3	5.0	6.7	6.7	5.0	3.3	1.7	(%)Lucky \hat{F}_γ^+	
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)		
Negative alpha \hat{T}_γ^- (%)	11.9	16.1	19.5	21.6	2.9	2.7	2.4	2.1	(%)Positive alpha \hat{T}_γ^+	
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)		
FDR ⁻	12.3	17.1	20.4	23.5	69.8	64.5	57.9	43.6	FDR ⁺	
Panel B: gross returns										
	Zero alpha ($\hat{\pi}_0$)				Negative alpha ($\hat{\pi}_A^-$)				Positive alpha ($\hat{\pi}_A^+$)	
Percentage	76.1				3.5				20.4	
	(0.02)				(0.01)				(0.01)	
Number of funds	2794				129				750	
Impact of Luck in the Left and Right Tails										
	Left Tail				Right Tail					
Signif.Level(γ)	0.05	0.10	0.15	0.20	0.20	0.15	0.10	0.05	Signif.Level(γ)	
Signif. \hat{S}_γ^- (%)	4.2	6.7	8.9	10.9	22.8	19.6	15.6	11.1	(%)Signif. \hat{S}_γ^+	
	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)		
Unlucky \hat{F}_γ^- (%)	1.9	3.8	5.7	7.6	7.6	5.7	3.8	1.9	(%)Lucky \hat{F}_γ^+	
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)		
Negative alpha \hat{T}_γ^- (%)	2.3	2.9	3.2	3.3	15.2	13.9	11.8	9.2	(%)Positive alpha \hat{T}_γ^+	
	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)		
FDR ⁻	45.7	57.0	63.9	69.5	33.4	29.1	24.3	17.1	FDR ⁺	

It is important to provide some explanation for the difference in estimates of the percentage of positive-alpha funds produced here compared to those of BSW. This is done with reference to Table 4. Following Cuthbertson et al. (2008)¹⁴, we document in Table 4 the results obtained from fund level regressions of different versions of 3F- and 4F-CAPM models discussed in Section 3. Estimates of parameters are obtained by averaging estimates produced from fund level regressions. We also report t-statistics associated with tests of significance for

¹⁴see Table 2 on page 620 of Cuthbertson et al. (2008)

these averaged estimates including those associated with each information variable.

Also reported there are results from testing for joint significance of the information variables on fund level regressions using the heteroskedastic robust Wald statistic. More specifically, two tests of this type are performed on the fund level regressions; PW1 records the proportion of Wald statistics that reject the joint null hypothesis that the lagged information variables had no effect on fund level parameters in equations (2) or (3). PW2 details the proportion of Wald-statistics that reject the joint null hypothesis that some of the lagged information variables do not affect the alpha in equation (3) given that some of these variables were found to affect the parameters in equation (2). The last three lines of this table report the proportion of *LM* tests that reject the null hypotheses of no first-, second- or third-order serial correlation in the errors.

Our first comment involves the *t*-statistics for the averaged estimated betas of the information variables. In aggregate, these variables do not affect the risk factors. In their studies Cuthbertson et al. (2008) and Cuthbertson et al. (2012), using UK investment trust data obtain similar results to us. Comparing the information measures, AIC (Akaike Information Criteria) and SBIC (Schwarz Bayesian Information Criteria), both indicate that the preferred model is the 4F-CAPM of Ferson and Schadt (1996) and not that of Carhart (1997). In this respect, our results are different from Cuthbertson et al. (2008) who find the unconditional model to have the smallest AIC and SBIC.

Results from tests of significance of the averaged estimated betas and alpha do not accurately reflect the importance of these variables in fund level regressions. For example, when the model of Ferson and Schadt (1996) is estimated on individual fund returns, the Wald-test rejects, in 57% of fund level regression, the joint null hypothesis that all parameters associated with the information variables are zero. So at least one of the lagged information variables affects the market risk factor in the majority of fund level regressions. When the model of Christopherson et al. (1998) is estimated using fund level returns, the Wald-statistic rejects, in 67.8% of the regressions, the joint null hypothesis that the lagged information variables affect alpha or the market risk factor. This test does not indicate the percentage of the regressions in which the information variables affect the alpha. To answer this question, an additional test was performed using the Wald statistic. This test found that, in 32.6% of fund level regressions where the information variables affect either alpha or the market risk factor, the information variables affected the alpha of these regressions.

The *LM* test for first-order serial correlation indicates that the proportion of regressions in which the residuals were correlated of first order ranged from 12.7% for the Christopherson et al. (1998) model to 14.7% for the unconditional 3F-CAPM. Tests for second-order serial correlation indicate that about 14.8% to 16% of the models had errors correlated of this order. A similar statement can be made regarding the errors of the regression models when tested for third-order correlation. These tests provide support for the use of the block bootstrap used here as well as by BSW, as the block bootstrap allows resampling without loss of serial

¹⁵ Many of the standard errors that are reported in the tables are zero because they are based on asymptotic theory of Genovese and Wasserman (2004) which requires the number of funds to go to infinity. Due to this theory, the larger the number of funds used the smaller the standard error. Because 3,673 funds were used, which is much larger than the 2076 funds used by Barras et al. (2010), many of the standard errors were 0 to two decimal places. Hence we report (0.00) in Tables 1, 2, 3, 6, 7 and 10 rather than report standard errors accurate to four decimal places.

correlation.

The fact that versions of the 4F-CAPM with time-varying parameters are more appropriate models of fund returns, as shown by the Wald-tests, is not without consequence for consistent estimation of the alpha and the percentages of negative, zero and positive-alpha mutual funds. In fact if one were to use the model of Carhart (1997) to estimate funds' alpha when this model is not appropriate for a large proportion of fund returns, this would lead to underestimation of funds' alpha and therefore an underestimation of these percentages. However, our approach to estimating the alphas is to use the encompassing model, either Ferson and Schadt (1996) only or Christopherson et al. (1998) only, for all funds to produce the results listed in Tables 2 or 3, even though one of these models will be the appropriate model of returns for only a large proportion but not all the funds in our sample. The encompassing model will nevertheless produce consistent estimates of fund alphas and also consistent estimates of these percentages.

Both asset pricing models model the relationship between betas and alpha as a reduced form. This is unlikely to hold for reasons mentioned in Section 3.4. In the next section, we determine whether a nonlinear version of the 4F-CAPM is more consistent with the mutual fund data used here rather than the linear versions estimated in this section.

Table 4:
Summary Statistics and Model Performance

The table provides the summary statistics for various factor models, including Fama-French three-factor model, Carhart's four factor model, conditional beta model, and conditional alpha-beta model. Reported statistics include estimates for fund's alpha, betas, adjusted R-squared, Schwartz Bayesian Information Criterion (SBIC) and Akaike Information Criterion (AIC). PW_1 records the proportion of Wald-tests of joint significance of the lagged informational variables (info1: 1-month T-Bill rate, info2: Dividend rate, info3: Term spread, info4: Default Spread) in the 4F-CAPM of Ferson and Schadt (1996). PW_2 records the proportion of Wald-tests of joint significance of lagged information variables that affect the alpha in the 4F-CAPM of Christopherson et al. (1998) given these variables affect the betas in the 4F-CAPM. Results from Lagrange Multiplier tests for first-, second- and third-order serial correlation of the errors are reported in the last three rows of this table. Each of these rows lists the proportion of the tests that reject the null hypothesis of no serial correlation of lag one ($LM.lag1$), lag two ($LM.lag2$) and lag three ($LM.lag3$) respectively. The sample consists of 3673 funds with at least 60 months of observations over the period from January 1984 to March 2015.

	3F-CAPM		4F-CAPM		Conditional Beta 3F-CAPM		Conditional Beta 4F-CAPM		Conditional Alpha-Beta 3F-CAPM		Conditional Alpha-Beta 4F-CAPM	
	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat
Alpha	-0.0008	-0.55	-0.0008	-0.58	-0.0005	-0.43	-0.0006	-0.42	-0.0015	-0.12	-0.0003	-0.05
$(R_m - R_f)$	0.9758	31.92	0.9785	31.30	0.8996	9.77	0.8794	9.63	0.8751	9.23	0.8669	9.24
SMB	0.2061	2.85	0.2054	2.93	0.2060	2.83	0.2052	2.91	0.2071	2.83	0.2065	2.90
HML	0.0049	0.84	0.0004	0.78	0.0016	0.62	0.0053	0.59	0.0076	0.75	0.0019	0.70
MOM			-0.0030	-0.10			-0.0040	-0.30			-0.0011	-0.12
info1*($R_m - R_f$)					-0.4738	-0.12	-2.7482	-0.06	-1.9045	-0.01	-3.5429	-0.04
info2*($R_m - R_f$)					-69.6511	-0.35	-73.1611	-0.43	-86.8971	-0.50	-80.8266	-0.51
info3*($R_m - R_f$)					1.7759	0.48	2.1262	0.56	1.4233	0.41	1.8130	0.48
info4*($R_m - R_f$)					-0.0043	-0.11	-0.0002	-0.14	-0.0108	-0.07	-0.0075	-0.10
info1									-0.2757	-0.15	-0.1199	-0.11
info2									1.1692	0.14	1.7175	0.16
info3									-0.0631	-0.33	-0.0792	-0.38
info4									0.0012	0.41	0.0009	0.32
Adj. Rsqd	0.8756		0.8836		0.8828		0.8898		0.8856		0.8920	
AIC	-1219.46		-1227.601		-1223.748		-1231.574		-1223.119		-1230.37	
SBIC	-1217.17		-1225.141		-1220.528		-1228.012		-1218.238		-1224.945	
PW_1					0.587		0.568		0.698		0.678	
PW_2									0.356		0.326	
$LM.lag1$	0.147		0.144		0.133		0.138		0.127		0.131	
$LM.lag2$	0.160		0.157		0.155		0.152		0.150		0.148	
$LM.lag3$	1.167		0.159		0.165		0.158		0.168		0.163	

7 Measuring fund performance using a nonlinear 4F-CAPM

Developing a nonlinear 4F-CAPM involves first testing for parameter changes in this model using the tests discussed in Section 3.4. Table 5 summarizes results from applying our CUSUM tests first to net and then to gross returns.

Our tests find a maximum of three changes in each of alpha or betas in the model of Carhart (1997) and no more than five changes are found in total in both alpha or betas. In Table 5, the first two blocks list the number of and proportion of changes in parameters of the 4F-CAPM detected by the CUSUM tests when the model is estimated using net returns. The third and fourth block also list the number and proportion of changes only when the model is alternatively estimated using gross returns. In the first application of the CUSUM tests to detect a change in the alpha or betas, we first estimate 3,673 regressions using net returns. The residuals from each estimated model are then used in the CUSUM tests to detect parameter changes. 302 of the fund level regressions are found to have a change in their alpha only, 204 are found to have a change in the betas only and 307 are found to have changes to both parameters. In the 2,860 remaining fund level regressions, our tests found no parameters had changed. For funds in which the tests found a change, the date of change is estimated and then the returns series separated into two parts; the first part consists of returns up to this date and the second part includes returns after it. To detect additional changes, the tests are applied to the residuals from models estimated on returns corresponding to one of the subparts.

In Table 5, the column headed "Forward tests" details the results from testing for additional changes after the date of break, while the column headed "Backward tests" applies the tests before this date. When the tests are applied forwards, a further 127 of the estimated models that had one change in alpha had a second change, 57 of these models had an additional change in betas and 45 have additional changes to both of them. Applying the CUSUM tests forwards and backwards for a second time, further changes in these parameters are found. Applying the tests both forwards and backwards for a third time found a small number of changes, while applying them forwards and backwards for a fourth time found no additional changes in parameters. We also find similar results when using gross rather than net returns with an additional 26 funds are found to have parameter changes in the 4F-CAPM that were not found using net returns. Such similarity in results from applying the CUSUM tests is to be expected as long as a fund does not alter their expense ratio.

Table 5: Summary of Results on Tests for Change Points in the Slopes and the Intercept: 1984-2015 (4F-CAPM)

	First test	Forward tests			Backward tests		
		1st	2nd	3rd	1st	2nd	3rd
Net returns							
	Number of funds						
Change in alpha & betas	307	45	4	0	12	3	0
Change in betas	204	57	14	1	36	2	0
Change in alpha	302	127	32	7	56	11	2
No change	2860	3444	3623	3665	3569	3657	3671
Percentage							
Change in alpha & betas	8.36	1.22	0.11	0.00	0.33	0.09	0.00
Change in betas	5.56	1.55	0.38	0.03	0.98	0.05	0.00
Change in alpha	8.22	3.46	0.87	0.19	1.52	0.30	0.05
No change	77.86	93.76	97.17	99.78	99.17	99.56	99.95
Gross returns							
	Number of funds						
Change in alpha & betas	317	52	5	0	15	2	0
Change in betas	202	52	12	1	37	3	0
Change in alpha	320	133	32	8	52	11	1
No change	2834	3436	3624	3664	3569	3657	3672
Percentage							
Change in alpha & beta	8.63	1.42	0.14	0	0.41	0.05	0
Change in beta	5.50	1.42	0.33	0.03	1.00	0.08	0
Change in alpha	8.72	3.62	0.87	0.22	1.42	0.30	0.03
No change	77.16	93.54	98.66	99.75	97.17	99.56	99.97

7.1 Net Returns

When net returns are used to estimate alpha, a fund that generates a positive alpha more than covers all management expenses and trading costs. Table 6 provides estimates of the percentage of negative, zero and positive-alpha funds obtained from the FDR using alpha estimated from the nonlinear 4F-CAPM where parameters are found to have changed. The estimated percentage of positive-alpha funds is now 8.4%, and statistically significant, 34.1% have a negative-alpha and 57.5% have a zero-alpha. In comparison to BSW's estimates, our estimate of positive-alpha funds is far greater than their small estimate of 0.6%. Our estimate of zero-alpha funds is much smaller than the 75% they find. They also find 24% of funds have a negative-alpha, a significantly smaller number than our finding of 57.5% we find. In sum, our estimates of negative, zero and positive-alpha funds derived from the nonlinear models are significantly different from estimates produced using the model of Carhart (1997) (cf. Table 1) which produces estimates of these percentages similar to those of BSW.

Table 6:

Percentages of Negative, Zero and Positive-Alpha Funds: 1984-2015 (Nonlinear 4F-CAPM)

Performance of funds is estimated using the 4F-CAPM over the entire sample period from January 1984 to March 2015. Provides estimated percentages of zero alpha ($\hat{\pi}_0$), negative alpha ($\hat{\pi}_A^-$) and positive alpha ($\hat{\pi}_A^+$) funds using the FDR approach based on the entire sample of 3673 funds and on net returns. This also includes significant funds ($\hat{S}_\gamma^-, \hat{S}_\gamma^+$) across the distribution of funds in the left tail and right tail using four different significance levels ($\gamma = 0.05, 0.1, 0.15, 0.2$). Furthermore, significant funds are decomposed into unlucky (\hat{F}_γ^-) and negative alpha (\hat{T}_γ^-) funds in the left tail, and lucky (\hat{F}_γ^+) and positive alpha (\hat{T}_γ^+) funds in the right tail. At the bottom of this table, we show the FDR in the left and right tail at different significance levels. Standard errors are indicated by parentheses and located below the estimate.

Percentage	Proportion of Negative Alpha and Positive Alpha Funds											
	Zero alpha ($\hat{\pi}_0$)				Negative alpha ($\hat{\pi}_A^-$)				Positive alpha ($\hat{\pi}_A^+$)			
	57.5 (0.01)				34.1 (0.02)				8.4 (0.01)			
	Impact of Luck in the Left and Right Tails											
	Left Tail				Right Tail							
Signif.Level(γ)	0.05	0.10	0.15	0.20	0.20	0.15	0.10	0.05	Signif.Level(γ)			
Signif. \hat{S}_γ^- (%)	19.0 (0.01)	24.5 (0.01)	29.1 (0.01)	32.6 (0.01)	14.2 (0.01)	12.5 (0.00)	10.5 (0.00)	7.9 (0.00)	(%)Signif. \hat{S}_γ^+			
Unlucky \hat{F}_γ^- (%)	1.4 (0.00)	2.9 (0.00)	4.3 (0.00)	5.7 (0.00)	5.7 (0.00)	4.3 (0.00)	2.9 (0.00)	1.4 (0.00)	(%)Lucky \hat{F}_γ^+			
Negative alpha \hat{T}_γ^- (%)	17.6 (0.01)	21.6 (0.01)	24.7 (0.01)	26.8 (0.01)	8.4 (0.01)	8.2 (0.01)	7.7 (0.00)	6.5 (0.00)	(%)Positive alpha \hat{T}_γ^+			
FDR ⁻	7.5	11.7	14.8	17.7	40.6	34.5	27.3	18.1	FDR ⁺			

As a robustness check to determine how sensitive our estimate of 8.4% is to different significance levels, we look at the percentages of significant funds in the left and right tails of the alpha distribution (\hat{S}_γ^- and \hat{S}_γ^+), percentages of unlucky and lucky (\hat{F}_γ^- and \hat{F}_γ^+) and percentages of truly negative-alpha or truly positive-alpha (\hat{T}_γ^- and \hat{T}_γ^+) funds. These estimates and how they change with the significance level can be found in Table 6. One would like estimates of the percentage of positive-alpha funds at each significant level to be similar because it shows that returns of positive-alpha funds are concentrated in the extreme right tail of the performance distribution and not distributed throughout the right tail¹⁶. The percentage of lucky funds ranges from a high of 5.7% to low of 1.4%, while the percentage of positive-alpha funds ranges from a high of 8.4% to a low of 6.5%. Each of our estimates is statistically significant whereas BSW find their estimates of truly positive-alpha funds to be statistically insignificant.

The reduction in the *FDR* follows from the large proportion of funds that report at least one change in their parameters. To see this, Table 5 reports that 8.36% of the funds experienced a change in alpha as well as the betas, while 5.56% and 8.22% reported changes in the betas or alpha respectively. This represents 22.14% of the funds which is a large percentage of the them. Using an estimate of the date that these parameters changed allowed for separate models to be estimated within each fund. According to this model, parameters can change in an arbitrary fashion and could be more accurately estimated. This produced larger *t*-statistics and led to smaller bootstrapped *p*-values. This reduced the number of false discoveries when

¹⁶For more explanation on this we refer to page 186 of BSW.

testing the null-hypothesis $H_0 : \alpha_{net,i} = 0$ versus $H_1 : \alpha_{net,i} > 0$ and therefore to a larger estimates in the percentages of zero- and positive-alpha funds. This led to the smaller $FDRs$ reported in Table 6 when compared the Tables 2 and 3, and also to the increase in these percentages.

7.2 Gross Returns

When gross returns are used to estimate the percentage of negative, zero and positive-alpha funds, interpretation of the results is that the fund is able to cover trading costs. Fund managers have little control over fund expenses because management companies set these with the approval of fund directors. Not subtracting fund management expenses from returns separates a fund's expense policy from the abnormal returns it generates. It is interesting, then, to re-examine these percentages by adding management expenses back to net returns producing gross returns. Here, we estimate the percentages of negative, zero and positive-alpha funds using the FDR technique but now returns are gross of management expenses. Table 7 details these estimates.

Table 7:

Percentages of Negative, Zero and Positive-Alpha Funds: 1984-2015 (Gross Returns, Nonlinear 4F-CAPM)

Performance of funds is measured using the 4F-CAPM over the entire sample period from January 1984 to March 2015. Estimated proportions of zero alpha ($\hat{\pi}_0$), negative alpha ($\hat{\pi}_A^-$) and positive alpha ($\hat{\pi}_A^+$) funds using the FDR approach based on the entire sample of 3673 funds and on gross returns. It also includes significant funds ($\hat{S}_\gamma^-, \hat{S}_\gamma^+$) across the distribution of funds in the left tail and right tail using four different significance levels ($\gamma = 0.05, 0.1, 0.15, 0.2$). Furthermore, significant funds are decomposed into unlucky (\hat{F}_γ^-) and negative alpha (\hat{T}_γ^-) funds in the left tail, and lucky (\hat{F}_γ^+) and positive alpha (\hat{T}_γ^+) funds in the right tail. At the bottom of this table, we show the FDR in the left and right tail at different significance levels. Standard errors are indicated by parentheses and located below the estimate.

Percentage	Zero alpha ($\hat{\pi}_0$)		Negative alpha ($\hat{\pi}_A^-$)		Positive alpha ($\hat{\pi}_A^+$)				
	68.3 (0.01)		11.9 (0.01)		19.8 (0.01)				
Impact of Luck in the Left and Right Tails									
	Left Tail				Right Tail				
Signif.Level(γ)	0.05	0.10	0.15	0.20	0.20	0.15	0.10	0.05	Signif.Level(γ)
Signif. $\hat{S}_\gamma^-(\%)$	9.0 (0.00)	12.3 (0.00)	15.1 (0.01)	17.6 (0.01)	24.6 (0.01)	21.8 (0.01)	18.7 (0.01)	14.5 (0.01)	(%)Signif. \hat{S}_γ^+
Unlucky $\hat{F}_\gamma^-(\%)$	1.7 (0.00)	3.4 (0.00)	5.1 (0.00)	6.8 (0.00)	6.8 (0.00)	5.1 (0.00)	3.4 (0.00)	1.7 (0.00)	(%)Lucky \hat{F}_γ^+
Negative alpha $\hat{T}_\gamma^-(\%)$	7.3 (0.00)	8.8 (0.00)	10.0 (0.01)	10.7 (0.01)	17.7 (0.01)	16.7 (0.01)	15.3 (0.01)	12.8 (0.01)	(%)Positive alpha \hat{T}_γ^+
FDR ⁻	19.0	27.8	33.9	38.9	27.8	23.5	18.2	11.8	FDR ⁺

Estimated percentages of negative, zero and positive-alpha funds reported in Table 7 change significantly compared to those produced on net returns using the same version of the 4F-CAPM. The estimated percentage of positive-alpha fund managers is 19.8% rather than 8.4% as it was for net returns. The percentage of negative-alpha funds is estimated to be 11.9%, whereas it was 34.1% for net returns. 68.3% of funds have zero-alpha which is much larger 57.8% estimated for net returns. Comparing results in Table 7 to estimates of

these percentages produced using the constant parameter 4F-CAPM, 12.6% (cf. Table 1) are positive-alpha funds, 7.1% and 80.3% (cf. Table 1) are estimated to be negative-alpha and zero-alpha funds respectively.

For the model of Christopherson et al. (1998), on the other hand, the estimates of these percentages, as reported in Table 3, are 20.4% positive-alpha, 3.5% negative-alpha and 76.1% zero-alpha funds respectively. For the most part, these percentages are similar across the three models. There is one interesting exception. The percentage of truly positive-alpha funds is estimated to be 25% for Ferson and Schadt (1996) rather than 20.4% for Christopherson et al. (1998). One would expect this because the model of Christopherson et al. (1998) allows the information variables to affect the alpha of this model. The effect of this on alpha would be to reduce its size because this version of the CAPM deducts from alpha any return that can be attributed to the use of publicly available information by the fund. BSW estimate these percentages as 9.6% positive-alpha, 4.5% negative-alpha, and 85.9% zero-alpha fund. Where BSW estimates differ from the corresponding estimates of the nonlinear model is in the percentage of positive alpha funds. Their estimate of positive-alpha funds is half as large as estimate produced using the models of Ferson and Schadt (1996), Christopherson et al. (1998) and the nonlinear 4F-CAPM models.

8 Performance of funds by investment styles

While we have been able to identify a somewhat higher proportion of funds with a positive-alpha than the existing studies. This begs the question as to whether the higher positive-alpha rate that we detect is related to the investment style of the Mutual Fund. To attract investors with different risk-return preferences, mutual funds follow different investment styles. The existing literature has come to differing views on the success rates of Mutual Funds. For example, Grinblatt and Titman (1989), Elton et al. (1996) show that aggressive growth and growth funds generate on average a positive-alpha and also have the highest percentage of positive-alpha funds. Hendricks et al. (1993), Daniel et al. (1997), Wermers (2000), Kosowski et al. (2006) and Huij and Verbeek (2007) support the view that both aggressive and growth funds have the highest percentage of positive-alpha funds. By contrast, Becker et al. (1999) argue that there is little difference in average returns between growth and income oriented funds. To investigate this matter in the light of our existing results, we estimate percentages of negative, zero and positive-alpha funds for the following three groups of funds: aggressive growth, growth and growth & income by using the false discovery rate technique.

Table 8 reports our estimates by investment style of the percentages of negative, zero and positive-alpha funds using the constant and unconditional 4F-CAPM on both net and gross returns. Looking at net returns using the constant parameters model, aggressive growth funds have the highest percentage of positive-alpha at 6.3%; for growth funds only 0.9% have positive alpha; and for growth & income funds we find 3.4% have positive alpha. However when we use the non-linear model there is a significant increase in the percentage of positive-alpha funds with 12.8% aggressive growth funds having positive alpha, 8.8% of growth funds having positive-alpha and 11% of growth & income having positive-alpha. Our estimates differ

significantly from the estimates of the existing literature; for example, when we use the model of Carhart (1997), 6.3% of aggressive growth funds are positive-alpha whereas BSW find only 3.9% are positive-alpha funds¹⁷. Our results confirm studies such as Hendricks et al. (1993) and Daniel et al. (1997) in showing that aggressive growth funds have the highest percentage of positive-alpha funds. From the results reported in Table 8, it should be noted that some funds change their style over their return history so funds can appear under two styles. This leads to a larger number of funds than those used to calculate the aggregate percentages reported in our earlier tables¹⁸.

When we look at gross returns, the percentages of positive and zero-alpha funds increases for all three categories of investment styles and consequently the percentage of negative-alpha funds falls. Growth & income funds have the highest percentage of positive-alpha at 22.1% followed by aggressive growth at 17.3% and growth funds at 16.1%. This may be partly explained by the fact that aggressive growth funds have substantially higher trading costs than the other two categories¹⁹ due to their trading activities. Consequently, on a gross returns basis, growth & income funds should have better performance than aggressive growth or growth-oriented funds.

¹⁷BSW in their analysis of different investment styles based on net returns find 75.5% of aggressive growth Funds are zero-alpha, 20.6% have a negative-alpha and 3.9% have a positive alpha.

¹⁸The difference number of funds (Table 8) between the constant model and nonlinear model is not due to the change in style because here funds are double counted for both models when they change styles. The larger number of funds for the nonlinear model compared to the number of funds used for to calculate the percentages for the constant parameter model arises because the change-point method (nonlinear model) when it locates a change-point breaks the funds in to two funds at the change-point so increasing the number of funds. The estimate of the parameter(s) that don't change is consistent but no longer efficient.

¹⁹See footnote Table 5 Grinblatt and Titman (1989).

Table 8:

Percentages of Negative, Zero, and Positive-alpha Funds: 1984-2015 (Different Investment Styles, Nonlinear vs. Time-invariant Parameters)

Performance of funds is measured using the 4F-CAPM with constant parameters and the nonlinear regression model with time-varying parameters, over the entire sample period from January 1984 to March 2015. The table shows the estimated proportions of zero alpha ($\hat{\pi}_0$), negative alpha ($\hat{\pi}_A^-$) and positive alpha ($\hat{\pi}_A^+$) funds on net and gross returns using the FDR approach for each category of funds, namely aggressive growth, growth, growth & income funds.

Investment styles	Net returns						Gross returns					
	Constant parameters			Nonlinear model			Constant parameters			Nonlinear model		
	$\hat{\pi}_0$	$\hat{\pi}_A^-$	$\hat{\pi}_A^+$	$\hat{\pi}_0$	$\hat{\pi}_A^-$	$\hat{\pi}_A^+$	$\hat{\pi}_0$	$\hat{\pi}_A^-$	$\hat{\pi}_A^+$	$\hat{\pi}_0$	$\hat{\pi}_A^-$	$\hat{\pi}_A^+$
Aggressive Growth												
Percentage	68.1	25.6	6.3	59.6	27.6	12.8	73.7	15.3	11.0	59.6	23.1	17.3
Number of funds	209	78	19	217	100	47	226	47	33	217	84	63
Growth												
Percentage	64.5	34.6	0.9	58.1	33.1	8.8	85.1	3.0	11.9	72.9	11.0	16.1
Number of funds	1586	851	21	1809	1030	274	2091	74	293	2287	315	505
Growth & Income												
Percentage	66.7	29.9	3.4	59.9	30.1	10.0	80.2	2.0	17.8	70.0	7.9	22.1
Number of funds	947	424	48	1060	533	177	1138	28	253	1247	140	394

9 Conclusions

In this paper we have provided up-to-date estimates of the percentage of positive-alpha producing funds according to the unconditional model of Carhart (1997). The model estimates only 0.8% of funds produce a positive-alpha which corresponds closely to the existing literature. However, when estimating the conditional models of Ferson and Schadt (1996) or Christopherson et al. (1998), we find in 66.5% and 67.2% respectively of fund level regressions that some of the estimates of the parameters associated with the information variables turn out to be statistically significant. Using conditional or our non-linear four-factor model, we estimate that between 2.9% and 8.4% of funds have positive-alpha on a net return basis. These estimates are significantly larger than the corresponding estimate of BSW and Fama and French (2010), who use the 4F-CAPM of Carhart (1997) and our own unconditional estimates of the percentage of positive-alpha producing funds. The estimated percentage of zero-alpha managers ranged from a low of 57.5% (Table 6) to a high of 67.2% (Table 1), while the percentage of negative-alpha funds ranged from a low of 29.9% (Table 2) to a high of 32.0% (Table 1).

We compare the three versions of the 4F-CAPM with our non-linear model which makes fewer assumptions than these models. The conditional model of Christopherson et al. (1998) nests the model of Ferson and Schadt (1996) which in turn nests the model of Carhart (1997). To develop our non-linear four factor model, novel CUSUM tests were used to detect possible changes in the parameters of the model of Carhart (1997). We found 22.14% of the funds in our sample experienced at least one change in either their alpha or the beta parameters; the model was then adapted to incorporate these changes. Our findings proved robust when using

nonlinear versions of the 3F-CAPM instead of nonlinear versions of the 4F-CAPM²⁰.

To provide further insights into our positive-alpha findings we also provide estimates of these percentages within different mutual fund investment styles. We find that aggressive growth funds have the highest percentage of positive-alpha funds at 12.8% on a net return basis, while 8.8% and 10.0% of growth and growth & income funds, respectively, produce positive-alpha. When returns are measured on a gross basis, this ordering changes. Growth & income funds, rather than aggressive growth funds, have the highest percentage of positive-alpha funds at 22.1%. This could be justified by the fact that aggressive growth funds aim to achieve the highest capital gains through active trading. Active trading implies high turnover and trading costs. Hence, for gross returns, aggressive growth funds have the lowest percentage of positive-alpha compared to growth and growth & income ones. The results are of direct relevance for investors who want to make informed decisions when choosing funds with different investment styles.

Our study contributes to the literature by showing that conditional models, especially ones that encompass a large subset of these models such as Christopherson et al. (1998), produce consistent estimates, on both a gross and net returns basis, of the proportions of negative-, zero- and positive-alpha funds should a nested model be more appropriate. Unlike Berk and van Binsbergen (2015), who use a different methodology than ours, we are still able to find some positive-alpha funds even after deduction of managerial expenses. However, we also agree with their finding that there are more negative-alpha funds than positive-alpha funds on both a gross and a net returns basis which suggests that investors need to be very careful when allocating their funds to active fund managers, and this can be partly done by looking at the investment style of the fund along with their suggestion of looking for persistence in performance.

Even though the estimates produced here of the proportion of actively managed mutual funds with positive alphas ranged from 2.9% to 8.4%, are larger than BSW, they still tell a cautionary tale. Both percentages are small and do not provide much hope to mutual fund investors, especially those saving for their retirement or looking for returns better than those offered by Exchange Traded Funds. Nevertheless, investors who wish to invest in actively managed mutual funds would be advised to restrict their search to funds having an aggressive growth style, as 12.8% are estimated to be positive-alpha funds on a net returns basis.

²⁰See Appendix A

Appendix A Robustness check

Table 9 calculates the percentages of negative-, zero-, and positive-alpha funds by estimating the alpha of the 4F-CAPM of Carhart (1997) using the 90 day T-bill rate rather than the 30 day T-bill rate. Comparing results reported here to those in Table 1, on a net returns bases, 67.2% (Table 1) of funds have zero-alpha compared to 66.7%, 32.0% (Table 1) have negative alpha compared to 32.4% and 0.8% (Table 1) are positive alpha compared to 0.9%. On a gross-returns basis, the results reported in these two tables are again very close. From this, we see that there is little change in the corresponding percentages regardless of the T-bill rate used to estimate this version of the 4F-CAPM. Results similar to this also apply when the 90 day T-bill rate is used in estimating these percentages via the conditional models. These tables are available from the authors upon request.

Table 9:

Percentages of Negative-, Zero- and Positive-Alpha Funds: 1984-2015 (Constant Parameter 4F-CAPM; Net and Gross Returns using 90 day treasury bill rate)

Performance of funds is estimated using the 4F-CAPM using equation (1) of Carhart (1997). Panel A (net returns) and Panel B (gross returns) show the estimated proportions of zero-alpha ($\hat{\pi}_0$), negative alpha ($\hat{\pi}_A^-$) and positive alpha ($\hat{\pi}_A^+$) funds using the FDR approach based on a dataset of 3,673 funds using the 4F-CAPM, for net returns and gross returns. Each panel also shows significant funds ($\hat{S}_\gamma^-, \hat{S}_\gamma^+$) across the distribution of funds in the left tail and right tail using four different significance levels ($\gamma = 0.05, 0.1, 0.15, 0.2$). Furthermore, significant funds are decomposed into unlucky (\hat{F}_γ^-) and negative alpha (\hat{T}_γ^-) funds in the left tail, and lucky (\hat{F}_γ^+) and positive alpha (\hat{T}_γ^+) funds in the right tail. At the bottom of each panel, we show the FDR in the left and right tail at different significance levels. The sample covers the period from January 1984 to March 2015. *Standard errors are indicated by parentheses and located below the estimate.*

Panel A: net returns, constant parameter model										
	Zero alpha ($\hat{\pi}_0$)				Negative alpha ($\hat{\pi}_A^-$)				Positive alpha ($\hat{\pi}_A^+$)	
Percentage	66.7 (0.02)				32.4 (0.02)				0.9 (0.00)	
Number of funds	2450				1190				33	
Impact of Luck in the Left and Right Tails										
	Left Tail				Right Tail					
Signif.Level(γ)	0.05	0.10	0.15	0.20	0.20	0.15	0.10	0.05	Signif.Level(γ)	
Signif. \hat{S}_γ^- (%)	13.4 (0.01)	19.3 (0.01)	24.2 (0.01)	28.1 (0.01)	7.4 (0.00)	5.8 (0.00)	4.4 (0.00)	2.6 (0.00)	(%)Signif. \hat{S}_γ^+	
Unlucky \hat{F}_γ^- (%)	1.7 (0.00)	3.3 (0.00)	5.0 (0.00)	6.7 (0.00)	6.7 (0.00)	5.0 (0.00)	3.3 (0.00)	1.7 (0.00)	(%)Lucky \hat{F}_γ^+	
Negative alpha \hat{T}_γ^- (%)	11.8 (0.01)	15.9 (0.01)	19.2 (0.01)	21.4 (0.01)	0.7 (0.01)	0.8 (0.00)	1.0 (0.00)	0.9 (0.00)	(%)Positive alpha \hat{T}_γ^+	
FDR ⁻	12.4	17.3	20.6	23.8	90.7	86.3	76.1	63.8	FDR ⁺	
Panel B: gross returns, constant parameter model										
	Zero alpha ($\hat{\pi}_0$)				Negative alpha ($\hat{\pi}_A^-$)				Positive alpha ($\hat{\pi}_A^+$)	
Percentage	75.1 (0.02)				10.5 (0.01)				14.4 (0.01)	
Number of funds	2758				384				531	
Impact of Luck in the Left and Right Tails										
	Left Tail				Right Tail					
Signif.Level(γ)	0.05	0.10	0.15	0.20	0.20	0.15	0.10	0.05	Signif.Level(γ)	
Signif. \hat{S}_γ^- (%)	3.9 (0.00)	6.4 (0.00)	8.6 (0.00)	11.0 (0.01)	20.4 (0.01)	17.3 (0.01)	13.8 (0.01)	8.7 (0.00)	(%)Signif. \hat{S}_γ^+	
Unlucky \hat{F}_γ^- (%)	1.9 (0.00)	3.8 (0.00)	5.6 (0.00)	7.5 (0.00)	7.5 (0.00)	5.6 (0.00)	3.8 (0.00)	1.9 (0.00)	(%)Lucky \hat{F}_γ^+	
Negative alpha \hat{T}_γ^- (%)	2.1 (0.00)	2.7 (0.00)	2.9 (0.01)	3.5 (0.01)	12.9 (0.01)	11.6 (0.01)	10.0 (0.01)	6.8 (0.00)	(%)Positive alpha \hat{T}_γ^+	
FDR ⁻	47.5	58.4	65.7	68.1	36.8	32.6	27.2	21.5	FDR ⁺	

To test whether our results are robust to different versions of the CAPM that are used to estimate fund alphas, we also estimate the percentage of negative, zero and positive-alpha funds based on the nonlinear version of the 3F-CAPM, the results are shown in Table 10. In general, the estimates of percentages of negative, zero and positive-alpha funds based on this model are similar to those produced using the linear 4F-CAPM. For instance, in terms of net returns, the estimated percentages are 9.1% positive-alpha, 33.0% negative-alpha and 57.9% zero-alpha respectively which is similar to those based on the 4F-CAPM: 8.4%, 34.1% and 57.5% reported in Table 6. When returns are measured on a gross basis we find 23.2% are positive-alpha, 12.4% negative-alpha, and 64.4% as zero-alpha. Estimates from the 3F-CAPM are similar to the percentages based on the 4F-CAPM confirming the latter's robustness.

Table 10:

Proportions of Negative, Zero and Positive-Alpha Funds: 1984-2015 (3F-CAPM, Nonlinear Model)

Performance of funds is measured using the nonlinear model based on the 3F-CAPM. Panel A shows the estimated proportions of zero alpha ($\hat{\pi}_0$), negative alpha ($\hat{\pi}_A^-$) and positive alpha ($\hat{\pi}_A^+$) funds using the FDR approach on net returns. The lower block of Panel A displays the detailed results including significant funds ($\hat{S}_\gamma^-, \hat{S}_\gamma^+$) across the distribution of funds in the left tail and right tail using four different significance levels ($\gamma = 0.05, 0.1, 0.15, 0.2$). Furthermore, significant funds are decomposed into unlucky (\hat{F}_γ^-) and negative alpha (\hat{T}_γ^-) funds in the left tail, and lucky (\hat{F}_γ^+) and positive alpha (\hat{T}_γ^+) funds in the right tail. The bottom of the Panel A shows the FDR for each significance level. Panel B shows results on gross returns. Standard errors are indicated by parentheses and located below the estimate.

Panel A: Net Returns											
	Zero alpha ($\hat{\pi}_0$)				Negative alpha ($\hat{\pi}_A^-$)				Positive alpha ($\hat{\pi}_A^+$)		
Percentage	57.9 (0.02)				33.0 (0.02)				9.1 (0.01)		
Impact of Luck in the Left and Right Tails											
	Left Tail				Right Tail						
Signif.Level(γ)	0.05	0.10	0.15	0.20	0.20	0.15	0.10	0.05	Signif.Level(γ)		
Signif. \hat{S}_γ^- (%)	19.6	25.0	29.1	32.2	14.9	13.4	11.5	8.9	(%)Signif. \hat{S}_γ^+		
	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.00			
Unlucky \hat{F}_γ^- (%)	1.4	2.9	4.3	5.8	5.8	4.3	2.9	1.4	(%)Lucky \hat{F}_γ^+		
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00			
Negative alpha \hat{T}_γ^- (%)	18.2	22.1	24.8	26.4	9.1	9.1	8.6	7.5	(%)Positive alpha \hat{T}_γ^+		
	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.00			
FDR ⁻	7.4	11.6	14.9	18.0	38.9	32.4	25.2	16.2	FDR ⁺		
Panel B: Gross Returns											
	Zero alpha ($\hat{\pi}_0$)				Negative alpha ($\hat{\pi}_A^-$)				Positive alpha ($\hat{\pi}_A^+$)		
Percentage	64.4 (0.02)				12.4				23.2		
Impact of Luck in the Left and Right Tails											
	Left Tail				Right Tail						
Signif.Level(γ)	0.05	0.10	0.15	0.20	0.20	0.15	0.10	0.05	Signif.Level(γ)		
Signif. \hat{S}_γ^- (%)	9.6	12.7	15.6	18.1	25.0	22.1	19.3	15.4	(%)Signif. \hat{S}_γ^+		
	0.00	0.00	0.01	0.01	0.01	0.01	0.01	0.01			
Unlucky \hat{F}_γ^- (%)	1.6	3.2	4.8	6.4	6.4	4.8	3.2	1.6	(%)Lucky \hat{F}_γ^+		
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00			
Negative alpha \hat{T}_γ^- (%)	8.0	9.5	10.8	11.7	18.5	17.2	16.1	13.8	(%)Positive alpha \hat{T}_γ^+		
	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01			
FDR ⁻	16.7	25.4	30.9	35.6	25.8	21.9	16.7	10.5	FDR ⁺		

References

- Admati, Anat R, Sudipto Bhattacharya, Paul Pfleiderer and Stephen A Ross (1986), 'On timing and selectivity', *The Journal of Finance* **41**(3), 715–730.
- Barras, Laurent, Olivier Scaillet and Russ Wermers (2010), 'False discoveries in mutual fund performance: Measuring luck in estimated alphas', *The Journal of Finance* **65**(1), 179–216.
- Becker, Connie, Wayne Ferson, David H Myers and Michael J Schill (1999), 'Conditional market timing with benchmark investors', *Journal of Financial Economics* **52**(1), 119–148.

- Berk, Jonathan B and Jules H van Binsbergen (2015), 'Measuring skill in the mutual fund industry', *Journal of Financial Economics* **118**(1), 1–20.
- Berk, Jonathan B and Richard C Green (2004), 'Mutual fund flows and performance in rational markets', *Journal of Political Economy* **112**(6), 1269–1295.
- Berndt, Ernst R (1991), *The practice of econometrics: classic and contemporary*, Addison-Wesley Reading, MA.
- Carhart, Mark M (1997), 'On persistence in mutual fund performance', *The Journal of Finance* **52**(1), 57–82.
- Christopherson, Jon A, Wayne E Ferson and Debra A Glassman (1998), 'Conditioning manager alphas on economic information: Another look at the persistence of performance', *Review of Financial Studies* **11**(1), 111–142.
- Cuthbertson, Keith, Dirk Nitzsche and Niall O'Sullivan (2008), 'UK mutual fund performance: Skill or luck?', *Journal of Empirical Finance* **15**(4), 613–634.
- Cuthbertson, Keith, Dirk Nitzsche and Niall O'Sullivan (2012), 'False discoveries in UK mutual fund performance', *European Financial Management* **18**(3), 444–463.
- Daniel, Kent, Mark Grinblatt, Sheridan Titman and Russ Wermers (1997), 'Measuring mutual fund performance with characteristic-based benchmarks', *The Journal of Finance* **52**(3), 1035–1058.
- Davidson, Russell and James G MacKinnon (2004), *Econometric theory and methods*, Vol. 5, Oxford University Press New York.
- Dybvig, Philip H and Stephen A Ross (1985), 'The analytics of performance measurement using a security market line', *The Journal of Finance* **40**(2), 401–416.
- Elton, Edwin J, Martin Jay Gruber and Christopher R Blake (1996), 'Survivor bias and mutual fund performance', *Review of Financial Studies* **9**(4), 1097–1120.
- Elton, Edwin J, Martin Jay Gruber, Sanjiv Das and Matthew Hlavka (1993), 'Efficiency with costly information: A reinterpretation of evidence from managed portfolios', *Review of Financial Studies* **6**(1), 1–22.
- Fama, Eugene F and Kenneth R French (2010), 'Luck versus skill in the cross-section of mutual fund returns', *The Journal of Finance* **65**(5), 1915–1947.
- Ferson, Wayne E and Rudi W Schadt (1996), 'Measuring fund strategy and performance in changing economic conditions', *The Journal of Finance* **51**(2), 425–461.
- Grinblatt, Mark and Sheridan Titman (1989), 'Mutual fund performance: An analysis of quarterly portfolio holdings', *Journal of Business* **62**(3), 393–416.
- Grinblatt, Mark and Sheridan Titman (1992), 'The persistence of mutual fund performance', *The Journal of Finance* **47**(5), 1977–1984.

- Grout, Paul and Anna Zalewska (2006), 'The impact of regulation on market risk', *Journal of Financial Economics* **80**(1), 149–184.
- Hendricks, Darryll, Jayendu Patel and Richard Zeckhauser (1993), 'Hot hands in mutual funds: Short-run persistence of relative performance, 1974-1988', *The Journal of Finance* **48**(1), 93–130.
- Henriksson, Roy D (1984), 'Market timing and mutual fund performance: An empirical investigation', *Journal of Business* **57**(1), 73–96.
- Huij, Joop and Marno Verbeek (2007), 'Cross-sectional learning and short-run persistence in mutual fund performance', *Journal of Banking & Finance* **31**(3), 973–997.
- Ippolito, Richard A (1989), 'Efficiency with costly information: A study of mutual fund performance, 1965-1984', *The Quarterly Journal of Economics* **104**(1), 1–23.
- Jensen, Michael C (1968), 'The performance of mutual funds in the period 1945–1964', *The Journal of Finance* **23**(2), 389–416.
- Kosowski, Robert, Allan Timmermann, Russ Wermers and Hal White (2006), 'Can mutual fund "stars" really pick stocks? New evidence from a bootstrap analysis', *The Journal of Finance* **61**(6), 2551–2595.
- Mamaysky, Harry, Matthew Spiegel and Hong Zhang (2008), 'Estimating the dynamics of mutual funds alphas and betas', *The Review of Financial Studies* **21**(1), 233–264.
- Pouliot, William (2016), 'Robust tests for change in intercept and slope in linear regression models with application to manager performance in the mutual fund industry', *Economic Modelling* **58**, 523–534.
- Storey, John D (2002), 'A direct approach to false discovery rates', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **64**(3), 479–498.
- Storey, John D, Jonathan E Taylor and David Siegmund (2004), 'Strong control, conservative point estimation and simultaneous conservative consistency of false discovery rates: a unified approach', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **66**(1), 187–205.
- The investment company fact book* (2015), 57 edn, Investment Company Institute.
- Vostrikova, L (1981), 'Detecting disorder in multidimensional random processes', *Soviet Mathematics Doklady* **24**, 55–59.
- Wermers, Russ (2000), 'Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses', *The Journal of Finance* **55**(4), 1655–1703.