

'Mutual Fund Performance: Skill or Luck ?'

Version : February 2005

MUTUAL FUND PERFORMANCE: SKILL OR LUCK?

Keith Cuthbertson*, Dirk Nitzsche*
and
Niall O'Sullivan**

Abstract:

Using a comprehensive data set on (surviving and non-surviving) UK mutual funds (April 1975 – December 2002), we use a bootstrap methodology to distinguish between 'skill' and 'luck' for *individual* funds. This methodology allows for non-normality in the idiosyncratic risks of the funds – a major issue when considering those funds which appear to be either very good performers or very bad performers, since these are the funds which investors are primarily interested in identifying. Our study points to the existence of genuine stock picking ability among a *small number* of top performing UK equity mutual funds (i.e. performance which is not solely due to good luck). At the negative end of the performance scale, our analysis strongly rejects the hypothesis that poor performing funds are merely unlucky. These funds demonstrate 'bad skill'.

Keywords : Mutual fund performance, Bootstrapping, Fama-French model

JEL Classification: C15, G11

* Cass Business School, City University, London

** Department of Economics, University College Cork, Ireland

Corresponding Author : Professor Keith Cuthbertson
Cass Business School, City University London
106 Bunhill Row, London, EC1Y 8TZ.
Tel. : 44-(0)-20-7040-5070 Fax : 44-(0)-20-7040-8881
E-mail : k.cuthbertson@city.ac.uk

1. We gratefully acknowledge the provision of mutual fund return data by Fenchurch Corporate Services using Standard & Poor's Analytical Software and Data.
2. Niall O' Sullivan is grateful for financial assistance from the Arts Faculty Research Fund at University College Cork.

MUTUAL FUND PERFORMANCE: SKILL OR LUCK?

Two key issues on fund performance have been central to recent academic and policy debates. The first is whether average risk adjusted abnormal fund performance (after expenses are taken into account) is positive, negative or zero. On balance, US studies of mutual (and pension) funds suggest little or no superior performance but somewhat stronger evidence of underperformance (e.g. Lakonishok et al 1992, Grinblatt, Titman and Wermers 1995, Daniel et al 1997, Carhart 1997, Chevalier and Ellison 1999, Wermers 2000, Pastor and Stamburgh 2002). Results using UK data on mutual and pension funds give similar results (e.g. Blake and Timmermann 1998, Blake, Lehmann and Timmerman 1999, Thomas and Tonks 2001).

A second major issue is whether abnormal performance can be identified ex-ante and for how long it persists in the future. Persistence is examined using either a contingency table approach or performance ranked portfolio strategies or by observing actual trades of mutual funds. Using the first two techniques the evidence is rather mixed. For US funds it seems that selecting funds with superior future performance is rather difficult and probably impossible, unless portfolio rebalancing is frequent (e.g. at least once per year) and the performance horizon is not longer than about +one-year (e.g. Grinblatt and Titman 1992, Hendricks, Patel and Zechauser 1993, Brown and Goetzmann 1995, Carhart 1997, Wermers 1997, Blake and Morey 2000, Bollen and Busse 2005, Mamaysky, Spiegel and Zhang 2004). A recent exception is Teo and Woo (2001) who find persistence in style adjusted returns for up to six years.

Studies using actual trades of mutual funds find that one-year persistence amongst winner funds is due to stocks passively carried over, rather than newly purchased stocks of winner funds performing better than newly purchased stocks of loser funds (Chen et al 1999). Following on from this Wermers (2003) finds that persistent large cash inflows to winner funds are invested with a lag and hence the average dollar invested in past winner funds does not earn more than that invested in past loser funds. This is consistent with the hypothesis of Berk and Green (2004) where excess fund returns are quickly bid away in a competitive market.

For UK data on mutual and pension funds there is little evidence of persistence in superior performance but much stronger evidence that poor performers continue to underperform (e.g. Blake and Timmermann 1998, Allen and Tan 1999, Fletcher and Forbes 2002, Blake, Lehmann and Timmermann 1999, Tonks 2004).

This study examines the performance of open-end mutual funds investing in UK equity (Unit Trusts and Open Ended Investment Companies OEICs) during the period April 1975 to

'Mutual Fund Performance: Skill or Luck ?'

December 2002. A data set of over 1,500 funds is examined. This represents almost the entire UK equity mutual fund industry at the end of the sample period. In comparison with the US mutual fund industry, there have been comparatively few studies of the performance of UK mutual funds (unit trusts). Unlike many previous studies the focus of this paper is on *individual* fund performance (particularly in the tails of the performance distribution) and in determining the role of luck versus skill.

In contrast to earlier studies which use 'conventional' statistical measures, often applied to portfolios of funds, we use a cross-section bootstrap procedure across *all* individual funds. This enables our 'luck distribution' for any chosen fund (e.g. the best fund), to encapsulate possible outcomes of luck not just for our chosen fund but across all the funds in our data set. We are then able to separate 'skill' from 'luck' in the performance of *individual* funds, even when the distribution of idiosyncratic risk across *many* funds is highly non-normal. This methodology has not been applied to UK data and was first applied to US mutual funds by Kosowski, Timmermann, White and Wermers (2004).

As noted above, the absolute performance of mutual (and pension) funds and the relative performance of active versus passive (index) funds are central to recent policy debates, particularly in Europe. With increasing longevity and given projected state pensions, a 'savings gap' is predicted for many European countries in 20 years time (Turner 2004, OECD 2003). Will voluntary saving in mutual and pension funds over the next 20 years be sufficient to fill this gap, so that those reaching retirement age have sufficient savings to provide an adequate standard of living? A key element here is the attractiveness of savings products in general and also the choice between actively managed and passive (or index/tracker) funds.

In recent theoretical and empirical work, the allocation across different asset classes (mainly bonds versus stocks, but in principal across all asset classes) has been examined in an intertemporal framework. The 'rule of thumb' that the percentage investment in risky assets (stocks) should equal '100 minus your age' is not robust either in the face of uncertain income (which gives rise to hedging demands – Bodie, Merton and Samuelson 1992, Campbell and Viceira 1999, Viceira 2001) or, when predictability is present (Brennan et al 1997, Viceira 1999) or, when there is uncertainty about parameters in the prediction equation (Barberis 2000, Xia 2001). In practice, the lack of a consensus 'model' of asset allocation at both the 'strategic' and 'tactical' level is starkly illustrated by Boots (the UK chemist) switching all its pension fund assets into bonds in 2001 (for strategic not market timing reasons), while most UK pension funds continue to hold around 70% in stocks. In the US, participants in 401(K) retirement plans (Benartzi and Thaler 2001), when faced with the choice between several funds each of which has alternative proportions of stocks and bonds, tend to use a simple 1/n allocation rule - so the

'Mutual Fund Performance: Skill or Luck ?'

actual allocation to each asset class is not determined by any sophisticated optimization problem and is changed infrequently. Such naïve asset allocation decisions may carry over to investment in mutual funds (and even trustees' decisions for pension fund asset allocations), so that poor funds survive and exacerbate the savings gap.

The behavioral finance literature (see Barberis and Thaler 2003 for a survey) has provided theoretical models and empirical evidence which suggests that active stock picking 'styles' such as value-growth (LaPorta et al 1997, Chan and Lakonishok 2004) and momentum (Jegadeesh and Titman 1993, 2001, Chan et al 2000), as well as market timing strategies (Pesaran and Timmermann 1994, 1995, 2000, Ang and Bekaert 2004) can earn abnormal returns after correcting for risk and transactions costs. Large sections of the mutual fund sector follow these active strategies and more recently there is an ongoing debate on whether mutual (and pension) funds should be allowed to invest in hedge funds and private equity, which also follow a wide variety of active strategies. The question is therefore whether one can find actively managed funds which outperform index funds (after correcting for risk and transactions costs).

The Presidential Commission on Social Security Reform (2001) and the State of the Union Address (2005) envisage the part-privatization of US Social Security. This will increase debate on all aspects of the fund management industry, particularly in the light of the 'market timing' abuses uncovered in the US by New York Attorney General Elliot Spitzer (Goetzmann, Ivkovic and Rouwenhorst (2001) - which has reduced confidence in the financial service sector's ability to provide adequate and fair treatment of retail investors. In the UK, the continuing switch from defined benefit to defined contribution pension schemes will strengthen the argument for a closer analysis of active versus passive strategies (as well as the competence and independence of trustee governance arrangements-Myners 2001).

The Financial Services Authority (FSA) in the UK is concerned that (retail) investors may be misled by mutual fund advertising. In its 'comparative tables' it currently does not enter a fund's ranking vis-a-vis competitor funds, in terms of (raw) returns. The FSA believes this could encourage more investment in funds which may simply have high returns because they are more risky (Blake and Timmermann 1998 and 2003 and Charles River Associates 2002).

To the extent that any 'savings gap' is to be filled by investment in mutual funds, the need to evaluate risk adjusted performance in a tractable and intuitive way, while taking account of the inherent uncertainty in performance measures, will be of increasing importance. This paper directly addresses the issue of 'skill versus luck'. We use 'alpha' α and the t-statistic of alpha t_{α} , as our measures of risk adjusted performance of mutual funds. However, we do not assume,

'Mutual Fund Performance: Skill or Luck ?'

as many earlier studies do, that a fund's idiosyncratic risk has a known parametric distribution. Instead we bootstrap the *empirical* distribution of idiosyncratic risk not just fund-by-fund, but across the whole cross-section of funds. This allows us to obtain a performance distribution for funds which are in the tails of the cross-section distribution – precisely the funds that investors are likely to be most interested in (i.e. extreme 'winners' or 'losers').

In fact, we mainly use t_α rather than 'alpha' α as our performance statistic since it has superior statistical properties. We also perform a number of bootstrap techniques to account for any serial correlation or heteroscedasticity in the idiosyncratic risk of each fund and possible contemporaneous cross-section correlation. The bootstrap procedure is robust to possible misspecification but reported results are of course dependent on the chosen performance model. We therefore examine a wide range of alternative models which we divide into three broad classes (i) unconditional models (Jensen 1968, Fama and French 1993, Carhart 1997) (ii) 'conditional-beta' models, in which factor loadings are allowed to change with conditioning public information (Ferson and Schadt 1996) and (iii) 'conditional alpha-beta' models where conditioning information also allows for time varying alphas (Christopherson, Ferson and Glassman 1998). We control for survivor bias by including 385 'nonsurviving' funds in the analysis and by examining the sensitivity of our bootstrap results to the inclusion of funds with different minimum 'data histories'.

We now anticipate some of our key findings. First the good news. The bootstrap procedure indicates there is strong evidence in support of genuine stock picking ability on the part of a relatively small number of 'top ranked' UK equity mutual funds. For example (using the Fama-French 3 factor unconditional model), of the top 20 ranked funds in the positive tail of the performance distribution, 11 funds exhibit levels of performance which cannot be attributable to 'luck' at 95% confidence, while 16 funds exhibit such performance at 90% confidence. As we move further towards the centre of the performance distribution (e.g. the top 20% to 50% of funds) these funds have positive alphas but this can be attributed to luck rather than skill.

In the left tail of the performance distribution, from the worst (ex-post) fund manager to the fund manager at the 40th percentile, we find that an economically significant negative abnormal performance cannot be attributed to bad luck but is due to 'bad skill'. Therefore there are a large number of truly appalling active funds in the universe of UK mutual funds. This is consistent with findings from the 'behavioral finance' literature where retail investors often use simple rules of thumb in asset allocation and who face learning and search costs when trying to evaluate alternatives.

'Mutual Fund Performance: Skill or Luck ?'

When examining different fund 'styles', we find genuine outperformance among the top equity income funds and equity growth funds but there is little evidence of skill for the top performers amongst the general equity and small stock funds. For all four investment styles the extreme right tail of the performance distribution indicates 'bad skill' rather than bad luck - which mirrors our earlier results when considering all funds together. We also find that the top ranked 'onshore funds' have genuine skill, whereas the positive alphas for the best 'offshore funds' are due to luck. For the left tails of these distributions, we find that extreme poor performers (negative alphas) demonstrate 'bad skill' rather than bad luck, whether they are onshore or offshore.

The above results are robust across all three classes of model we investigate, across several variants of the bootstrap and do not appear to be subject to survivor bias. The strong message from these results is that there are a few 'top funds' who have genuine skill but the majority have either no skill and do well because of luck or perform worse than bad luck and essentially waste investors time and money. If you choose your active funds by throwing darts at the *Financial Times* or *Wall Street Journal* mutual fund pages, then you are highly likely to choose a fund which has no skill - you would be better off choosing an index fund (especially after transactions costs). On the other hand a careful analysis of risk adjusted performance taking full account of luck across *all* funds, can identify, with reasonable probability, those few funds with genuine skill.

In the rest of the paper we proceed as follows. Section 1 describes the data used in the study. In section 2 we discuss performance measurement models applied to mutual fund returns. Section 3 details the bootstrap methodology. In section 4 we evaluate the performance measurement models and select a subset of 'best models' to which we apply the bootstrap procedure. Section 5 examines the results of the bootstrap analysis and section 6 concludes.

1. Data

Our mutual fund data set comprises 1,596 equity Unit Trusts and Open Ended Investment Companies (OEICs). These funds invest primarily in UK equity (i.e. minimum 80% must be in UK equities) and represent almost the entire set of equity funds which have existed at any point during the sample period under consideration, April 1975 – December 2002. Unit trusts are 'open ended' mutual funds and they can only be traded between the investor and the trust manager, there is no secondary market. They differ from 'investment trusts' which are closed end funds. Mutual fund monthly returns data have been obtained from Fenchurch Corporate Services using Standard & Poor's Analytical Software and Data. By restricting funds to those investing in

'Mutual Fund Performance: Skill or Luck ?'

UK equity, more accurate benchmark factor portfolios may be used in estimating risk adjusted abnormal performance.

Among the database of 1,596 funds, a total of 334 funds are referred to as 'second units'. These arise because of mergers or 'splits' and in the vast majority of cases the mergers occur early and the splits occur late in the fund's life, and therefore these second units report relatively few 'independent' returns. We exclude 334 such funds from the analysis. Furthermore, 112 of the funds in the database are market (FTSE 250) index/tracker funds and as we are interested in stock selection ability, these are also excluded. This leaves 1,150 non-tracker independent (i.e. non-second unit) funds for our analysis.

The equity funds are categorized by the investment objectives of the funds which include: equity growth (122 funds), equity income (253 funds), general equity (i.e. income and growth, 591 funds) and smaller companies (184 funds). The data set includes both surviving funds (765) and nonsurviving funds (385). Nonsurviving funds may cease to exist because they were merged with other funds or they may have been forced to close due to bad performance. Because of the latter scenario, it is critical to include nonsurviving funds in any performance analysis of the mutual fund industry, as failure to do so may bias performance findings upwards (Carhart et al 2002). However, it would be misleading to assume that all such funds died due to poor performance (Blake and Timmermann 1998), since some mergers may arise due to good rather than bad performance.

In addition, funds are also categorized by the location of operation. Onshore funds (845) are managed in the UK while offshore funds (305) are operated from Dublin, Luxembourg, Denmark, the Channel Islands or some other European locations.

All fund returns are measured gross of taxes on dividends and capital gains. The bid/offer spread in fund returns captures the initial charge (typically 5%), stamp duty, dealing charges and commissions incurred by the fund, and the bid/offer spreads of the underlying securities. However, because our focus is on the performance of the fund's managers rather than on returns to investors/customers, the returns data in this study are calculated bid-price to bid-price (with income reinvested), and therefore exclude such transaction costs.

The market factor used is the FT All Share Index of total returns (i.e. including dividends). Excess returns are calculated using the one-month UK T-bill rate. The factor mimicking portfolio for the size effect, SMB, is the difference between the monthly returns on the Hoare Govett Small

'Mutual Fund Performance: Skill or Luck ?'

Companies (HGSC) Index and the returns on the FT 100 index¹. The value premium, HML, is the difference between the monthly returns of the Morgan Stanley Capital International (MSCI) UK value index and the returns on the MSCI UK growth index². The factor mimicking portfolio's one-year momentum behavior, PR1YR, has been constructed using the FT100 (total return) index³.

Other variables used in conditional and market timing models include the one-month UK T-bill rate, the dividend yield on the FT-All Share index and the slope of the term structure (i.e. the yield on the UK 20 year gilt minus the yield on the UK three-month T-bill).

2. Performance Models

The alternative models of performance we consider are well known 'factor models' and therefore we only describe these briefly. Each model can be represented in its unconditional, conditional-beta and conditional alpha-beta form. For all models the intercept ('alpha') α and in particular the t-statistic of alpha t_α , are our measures of risk adjusted abnormal performance.

Unconditional Models

These have factor loadings that are time invariant. The alpha α_i of the CAPM or market model (Jensen 1968) is given by the regression:

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

where $r_{i,t} = (R_{i,t} - R_{f,t})$, $R_{i,t}$ = expected return on fund-i in period t , $R_{f,t}$ = risk free rate, $r_{m,t} = (R_{m,t} - R_{f,t})$ is the expected excess return on the market portfolio.

Carhart's (1997) performance measure is the alpha estimate from a four-factor model:

¹ The HGSC index measures the performance of the lowest 10% of stocks by market capitalization, of the main UK equity market. Both indices are total return measures.

² These indices are constructed by Morgan Stanley who ranks all the stocks in their UK national index by their book-to-market ratio. Starting with the highest book-to-market ratio stocks, these are attributed to the value index until 50% of the market capitalization of the national index is reached. The remaining stocks are attributed to the growth index. The MSCI national indices have a market coverage of at least 60% (more recently this has been increased to 85%). Total return indices are used for the construction of the HML variable.

³ For each month, the equally weighted average returns of stocks with the highest and lowest 30% returns over the previous 11 months are calculated. The PR1YR variable is then constructed by taking the difference between these two variables.

'Mutual Fund Performance: Skill or Luck ?'

$$(2) \quad r_{i,t} = \alpha_i + \beta_{1i} r_{m,t} + \beta_{2i} SMB_t + \beta_{3i} HML_t + \beta_{4i} PRIYR_t + \varepsilon_{i,t}$$

where SMB_t , HML_t and $PRIYR_t$ are factor mimicking portfolios for size, book-to-market value and one-year momentum effects, respectively. On US data, Fama and French (1993) find that a three-factor model including $r_{m,t}$, SMB_t and HML_t factors, provides significantly greater power than the CAPM alone and Carhart (1997) finds that $PRIYR_t$ is statistically significant in explaining (decile) returns on US mutual funds.

Conditional-Beta Models

Conditional models (Ferson and Schadt 1996) allow for the possibility that a fund's factor betas depend on lagged public information variables. This may arise because of under and overpricing (Chan 1988 and Ball and Kothari 1989), or changing financial characteristics of companies such as gearing, earnings variability and dividend policy (Foster 1986, Mandelker and Rhee 1984, Hochman 1983, Bildersee 1975). Also, an active fund manager may alter portfolio weights and consequently portfolio betas depending on public information. Thus there may well be time variation in the portfolio betas depending on the information set Z_t so that $\beta_{i,t} = b_{0i} + B_2'(z_t)$, where z_t is the vector of deviations of Z_t from its unconditional mean. For the CAPM this gives:

$$(3) \quad r_{i,t+1} = \alpha_i + b_{0i}(r_{b,t+1}) + B_i'(z_t * r_{b,t+1}) + \varepsilon_{i,t+1}$$

where $r_{b,t+1}$ = the excess return on a benchmark portfolio (i.e. market portfolio in this case). The null hypothesis of zero abnormal performance is $H_0: \alpha_i = 0$. The conditional-beta approach can also be applied to the Fama-French and Carhart models (although the number of parameters to estimate can rise quite quickly).

Conditional Alpha-Beta Models

Christopherson, Ferson and Glassman (1998) assume that alpha (as well as the beta's) may depend linearly on z_t so that $\alpha_{i,t} = \alpha_{0i} + A_i'(z_t)$ and the performance model is:

$$(4) \quad r_{i,t+1} = \alpha_{0i} + A_i'(z_t) + b_{0i}(r_{b,t+1}) + B_i'(z_t * r_{b,t+1}) + \varepsilon_{i,t+1}$$

'Mutual Fund Performance: Skill or Luck ?'

Here, α_{0i} measures abnormal performance after controlling for (i) publicly available information, z_t and (ii) adjustment of the factor loadings based on publicly available information.

Following earlier studies (Ferson and Schadt 1996, Christopherson, Ferson and Glassman 1998) our Z_t variables include permutations of : the one-month T-Bill yield, the dividend yield of the market factor and the term spread.

Market Timing

In addition to stock selection skills, models of portfolio performance also attempt to identify whether fund managers have the ability to market-time. Can fund managers successfully assess the future direction of the market in aggregate and alter the market beta accordingly ? Suppose a successful market timer receives a private signal regarding the future direction of the market and adjusts the market factor loading $\beta_{it} = \theta_i + \gamma_{im}[r_{m,t}]$ so that (1) may now be written (see Treynor and Mazuy 1966):

$$(5) \quad r_{i,t} = \alpha_i + \theta_i(r_{m,t}) + \gamma_{im}[r_{m,t}]^2 + \varepsilon_{i,t}$$

where $\gamma_{im} > 0$ is the unconditional measure of market timing ability. Alternatively, the Merton and Henriksson (1981) model of market timing is:

$$(6) \quad r_{i,t} = \alpha_i + \theta_i(r_{m,t}) + \gamma_{im}[r_{m,t}]^+ + \varepsilon_{i,t}$$

where $[r_{m,t}]^+ = \max\{0, r_{m,t}\}$ and γ_{im} is the unconditional measure of market timing ability. These two models can be easily generalized to a conditional-beta model, where β_i also depends on the public information set, z_t (Ferson and Schadt 1996).

As a test of robustness, each of the above models is estimated for each mutual fund. Results are then averaged across funds in order to select a single 'best fit' model from each of the three classes : unconditional, conditional-beta and conditional alpha-beta models. These three 'best' models are used in the subsequent (computationally intensive) bootstrap analysis.

3. Bootstrap Methodology

'Mutual Fund Performance: Skill or Luck ?'

Previous studies of UK unit trust performance all use 'conventional' statistical measures, and generally find (using a three or four factor model) that there is little or no positive abnormal performance by (portfolios of) 'best' funds, whereas the 'worst' funds have statistically significant negative risk adjusted performance (see *inter alia*, Blake and Timmermann 1998, Quigley and Siquefield 2000, Fletcher and Forbes 2002). Among US mutual funds there is little evidence of positive abnormal performance but stronger evidence of poor performing funds - Carhart (1997), Christopherson et al (1998), Hendricks et al (1993). It has been argued that abnormal performance may be due to a momentum effect in the stock holdings rather than genuine stock picking skill (Carhart 1997, Chen et al 2000), although the evidence is not entirely definitive (Chen et al 2000 and Wermers 2000).

In this paper we use a cross-section bootstrap procedure and are able to separate 'skill' from 'luck' for *individual* funds, even when idiosyncratic risks are highly non-normal – as is the case for funds in the extreme tails, in which investors are particularly interested. We begin with a largely intuitive exposition of our bootstrap analysis, using 'alpha' as our measure of risk adjusted abnormal performance and the market model (CAPM) as the 'true model' of expected fund returns. In a large universe of funds (say $n = 1,000$) there will always be some funds that perform well (badly), simply due to good (bad) luck.

Assume that when *all* funds have no stock picking ability (i.e. $\alpha_i = 0$ for $i = 1, 2, \dots, n$) each fund's 'true' alpha is normally distributed and each fund has a different but *known* standard deviation of σ_i . Suppose we are interested in the performance of the best fund. If we 'replay history' just for the 'best fund', where we impose $\alpha_i = 0$ (here $i = \text{best fund}$), but 'luck' is represented by the normal distribution with known standard deviation σ_i , we would sample a different estimate of alpha. Of course there is a high probability that we sample a value of alpha close to zero, but 'luck' implies that we may sample a value for alpha which is in the extreme tails of the distribution. Similarly, when we resample the alpha for all the other $n-1$ funds, with all $\alpha_i = 0$ (but with different σ_i), it is quite conceivable that the second or third etc ranked fund in the *ex-post* data now has the highest alpha. This would hold a fortiori if the distributions of the second, or third, etc. ranked funds have relatively large values of σ_i .

From this single 'replay of history' where we impose $\alpha_i = 0$ across all funds, we have $(\alpha_1^{(1)}, \alpha_2^{(1)}, \dots, \alpha_n^{(1)})$ from which we choose the largest value $\alpha_{\max}^{(1)}$. So taking the 'luck distribution' across *all funds* into consideration (with different σ_i 's), we now have one value

'Mutual Fund Performance: Skill or Luck ?'

$\alpha_{\max}^{(1)}$ for the best fund which arises purely due to sampling variability or luck. However, by repeating the above (B-times) and each time choosing $\alpha_{\max}^{(k)}$ (for $k = 1, 2, \dots, B$ trials) we can obtain the complete distribution of α_{\max} , under the null of no outperformance. This 'empirical' distribution for α_{\max} is purely due to luck.

Note that the distribution of α_{\max} uses the information about 'luck' represented by all the funds and not just the 'luck' encountered by the 'best fund' in the ex-post ranking. This is a key difference between our study and many earlier studies that have used this type of methodology. It is important to measure the performance distribution of the 'best fund' not just by re-sampling from the distribution of the best fund ex-post, since this is a single realization of 'luck' for one particular fund. Clearly, re-running history for just the ex-post best fund ignores the other possible distributions of luck (here just the different standard deviations) encountered by all other funds – these other 'luck distributions' provide highly valuable and relevant information.

Having obtained our 'luck distribution', we now compare the best fund's actual *ex-post* performance given by its estimated $\hat{\alpha}_i$ against the 'luck distribution' for the best fund. If the actual performance ($\hat{\alpha}_i$) exceeds the 5% cut off point in the luck distribution we can reject the null hypothesis (at 95% significance), that the performance is attributable to luck.

Out of our 1,000 actual funds we have only one best fund, based on its measured ex-post alpha over say a 20 year span of (monthly returns) data. Suppose the ex-post alpha of the best fund is 1.7% per month (20.4% p.a.). This looks like an impressive performance but is it within the bounds set by luck or, has the best fund really demonstrated genuine stock picking ability? Above we simulated the complete distribution of α_{\max} under the null of no outperformance and therefore we know the 5% right tail cut off point which is say 2% p.m. Therefore at a 95% confidence level we cannot reject the null hypothesis that the performance of the 'actual ex-post best fund' is purely due to luck. Of course had the ex-post alpha of the best fund been say 2.3% p.m. then we can reject the null and infer that the ex-post best fund really does have genuine stock picking skills (at better than 95% confidence).

Above, we could have chosen any fund (e.g. the 2nd best fund) on which to base the 'luck distribution'. So, we can compare the actual *ex-post* ranking for any chosen fund against *its* luck distribution and separate luck from skill, for all *individual* funds in our sample.

'Mutual Fund Performance: Skill or Luck ?'

The above demonstrates the main features and intuition behind our analysis of fund performance used in this paper. But a key difference in our study (which we highlight below) is that under the null of no out-performance, we do *not assume* the distribution of alpha for each fund is normal and *each fund's alpha* can in principal take on any distribution. The distribution for each fund's 'luck' is represented by the empirical distribution observed in the historic data and this distribution can be different for each fund. Hence the distribution of α_{\max} (under the null) encapsulates all of the different individual fund's 'luck distributions'.

Investors are particularly interested in funds in the tails of the performance distribution, such as the best fund, the second best fund, and so on. We find that the 'luck distribution' of alpha for these funds are highly non-normal, thus invalidating the usual test statistics. This motivates the use of the cross section bootstrap to ascertain whether the 'outstanding' or 'abysmal' performance of 'tail funds' is due to either, good or bad skill or good or bad luck, respectively.

There are a number of possible explanations as to why non-normal security returns can remain at the portfolio (mutual fund) level. As noted by Kosowski et al (2004), co-skewness of individual constituent non-normal security returns may not be diversified away in a fund⁴. Also, funds may hold derivatives to hedge return outcomes and this may result in a non-normal return distribution.

Basic Bootstrap

Kosowski et al (2004) provide a thorough analysis of the bootstrap methodology applied to mutual fund performance so we provide only a brief exposition of the basic procedure (Politis and Romano 1994). Consider an estimated model of equilibrium returns of the form:

$$(7) \quad r_{i,t} = \hat{\alpha}_i + \hat{\beta}_i' X_t + e_{i,t}$$

for $i = \{1, 2, \dots, n\}$ funds, where $T_i =$ number of observations on fund- i , $r_{i,t} = (R_{i,t} - r_{f,t})$, $X_t =$ matrix of risk factors and $e_{i,t} =$ residuals of fund- i . For our 'basic bootstrap' we use residual-only resampling, under the null of no outperformance. This involves the following steps (Efron and Tibshirani 1993). First, estimate the chosen model for each fund (separately) and save the $\{\hat{\beta}_i, e_{i,t}\}$. Next, for each fund- i , draw a random sample (with replacement) of length T_i from the

⁴ The central limit theorem implies that a large, well diversified and equal weighted portfolio of non-normally distributed securities will approximate normality. However, many funds do not have these characteristics.

'Mutual Fund Performance: Skill or Luck ?'

residuals $e_{i,t}$. While retaining the original chronological ordering of X_t , use these *re-sampled* bootstrap residuals $\tilde{e}_{i,t}$ to generate a simulated excess return series $\tilde{r}_{i,t}$ for fund-i, under the null hypothesis of no abnormal performance (i.e. setting $\alpha_i = 0$):

$$(8) \quad \tilde{r}_{i,t} = 0 + \hat{\beta}_i' X_t + \tilde{e}_{i,t}$$

By construction, the 'true' abnormal performance, for fund-i is zero. This is then repeated for all funds. Next, using the simulated returns $\tilde{r}_{i,t}$, the performance model is estimated and the resulting estimate of alpha $\tilde{\alpha}_i^{(1)}$ for each fund is obtained. The $\tilde{\alpha}_i^{(1)}$ estimates for each of the n-funds represent sampling variation around a true value of zero (by construction) and are entirely due to 'luck'. The $\tilde{\alpha}_i^{(1)}$ $\{i = 1, 2, \dots, n\}$ are then ordered from highest to lowest. The above process is repeated B times for each of the n funds, where B (= 1,000) denotes the number of bootstrap simulations. The bootstrap estimates of $\tilde{\alpha}_i$ may be gathered in a matrix of dimension (n x B) as follows.

$$\begin{bmatrix} \tilde{\alpha}_1^{(1)} & \tilde{\alpha}_1^{(2)} & \dots & \tilde{\alpha}_1^{(B)} \\ \tilde{\alpha}_2^{(1)} & \tilde{\alpha}_2^{(2)} & \dots & \tilde{\alpha}_2^{(B)} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\alpha}_n^{(1)} & \tilde{\alpha}_n^{(2)} & \dots & \tilde{\alpha}_n^{(B)} \end{bmatrix}$$

The first row of this *sorted* 'bootstrap matrix' now contains the highest values of $\tilde{\alpha}_i$ from the B bootstrap simulations, under the null hypothesis $\alpha_i = 0$. This is the 'luck distribution' for the extreme top performer. The second row contains the second highest values of $\tilde{\alpha}_i$ etc. Therefore each row of the bootstrap matrix provides a separate distribution of performance, for *each* point in the performance distribution, from the extreme best performer to the extreme worst performer, all of which are solely due to luck.

We can now compare any alpha from the performance model using *actual* data, with its appropriate 'luck distribution' Suppose we are interested in whether the performance of the ex-post best fund $\hat{\alpha}_{\max}$ is due to skill or luck. If $\hat{\alpha}_{\max}$ is greater than the 5% upper tail cut off point for the 'luck distribution' for the highest ranked fund (i.e. the first row of the bootstrap matrix) then we reject the null that its performance is due to luck (at 95% confidence). We infer that the fund

'Mutual Fund Performance: Skill or Luck ?'

has genuine skill. This can be repeated for any other point in the performance distribution, right down to the ex-post worst performing fund in the data (which is compared against the bootstrap distribution of worst possible performance – the last row of the bootstrap matrix). The latter shows whether the worst performing fund, does worse than can be explained by bad luck alone.

However, notwithstanding the above exposition in terms of the 'luck distribution' for alpha, our bootstrap analysis mainly focuses on the 'luck distribution' for the *t*-statistic of alpha $t_{\tilde{\alpha}_i}$ because it has better statistical properties (i.e. it is a pivotal statistic, see Kosowski et al 2004 and Hall 1992, for further discussion). The intuitive reason for this is straightforward. The idiosyncratic risk of funds with few observations may have high variance and will in consequence tend to generate 'outlier alphas'. These funds may disproportionately occupy the extreme tails of the bootstrapped alpha distributions leading to a very high variance in their 'luck distribution'. However, $t_{\tilde{\alpha}_i}$, scales alpha by its estimated standard error and therefore is independent of the 'nuisance parameter' $\sigma_{\varepsilon_i}^2$ and has superior statistical properties. The argument applies *a fortiori* at the extremes of the performance distribution – which are of particular interest.

Throughout this study both the ex-post actual and bootstrap *t*-statistics are based on Newey-West heteroscedasticity and autocorrelation adjusted standard errors. In our baseline bootstrap we set the minimum number of observations for the inclusion of any fund in the analysis at $T_{i,\min} = 60$ months to further improve statistical reliability.

4. Model Selection

In this section, the equilibrium models described in section 2 are estimated. For each model, cross-sectional (across funds) average statistics are presented. Based on these statistics, a single 'best model' is chosen from each of the 3 model classes; (i) unconditional, (ii) conditional-beta and (iii) conditional alpha-beta. All tests are conducted at a 5% significance level unless stated otherwise and results presented relate to all UK equity mutual funds over the period April 1975 – December 2002 and are based on funds with $T_{i,\min} = 60$.

Table 1, reports key statistics for a sub-set of models within each class : panel A, (unconditional), panel B (conditional-beta) and panel C (conditional alpha-beta). The cross-sectional average alpha takes on a small and statistically insignificant negative value for most models. (The exceptions being for some of the conditional alpha-beta models (Panel C) and the market timing models ('TM' and 'HM' – models 4, 5 and 10), where the average alpha is positive

'Mutual Fund Performance: Skill or Luck ?'

but again not statistically significant). The finding of negative abnormal performance (on average) is consistent with Blake and Timmermann (1998). They report results for equally weighted portfolios of UK mutual funds, which are in line with our results in the bottom half of panels A and B, where we find that the alpha of an *equally weighted* portfolio of all funds, (for all Fama-French and Carhart models) have statistically significant negative alphas.

[Table 1 here]

However, of key importance for this study (and for investors) is the relatively large cross-sectional standard deviations of the alpha estimates which is around 0.25% p.m. (3% p.a.), for most of the models in panels A and B and slightly larger for conditional alpha-beta models in panel C. This implies that the extreme tails of the distribution of abnormal performance may contain a substantial number of funds. This is important since investors are more interested in funds in the right tail of the performance distribution and avoiding those in the extreme left tail, than they are in the 'average fund's' performance.

The market excess return, $r_{m,t}$, and the *SMB* factor betas are consistently found to be statistically significant across all models, whereas the *HML* factor beta is often only just statistically significant (at a 10% significance level). The momentum factor (*PR1YR*) is also rather poorly determined, in contrast to US studies (Carhart 1997). On the basis of estimated betas the Fama-French 3 factor model appears the 'best' unconditional model (Panel A, 'model 2'). For conditional-beta models (panel B) only the dividend yield variable appears to produce near statistically significant results for a Fama-French 3 factor model. In the conditional alpha-beta model none of the conditional alphas has a t-statistic greater than 1.11 (Panel C : the Z1, Z2, Z3 variables) but some of the conditional betas are bordering on statistical significance (in particular those for the Fama-French 3 factor model – model 15). The Treynor-Mazuy and Merton-Henriksson market timing variables are not statistically significant.

The above results suggest that the unconditional Fama-French 3 factor model explains the UK mutual fund performance data reasonably well (Panel A, model 2). There is little additional explanatory power from the conditional and market timing variables. This finding is consistent with existing studies of UK market timing (Fletcher 1995, Leger 1997) while Jiang (2003) also finds against superior market timing using nonparametric tests on US equity mutual funds.

Turning now to diagnostics (bottom half of Panels A, B and C), the adjusted R^2 across all models is around 0.8, while around 70% of funds have non-normal errors (Bera-Jarque statistic),

'Mutual Fund Performance: Skill or Luck ?'

and around 40% of funds have serial correlation (which is of order one – LM statistic). The Schwartz information criterion (SIC) does not vary greatly across all model specifications but indicates that the best unconditional and conditional-beta model is the Fama-French 3 factor model (panel A, 'model 2' and panel B, 'model 8' respectively, where the conditional-beta model only includes the dividend yield variable).

Although model 14 (Panel C) has $SIC = 1.331$ it is very similar to model 8 above, so in the interests of presenting results from diverse specifications, 'model 15 – FF(3)' is taken as the 'best' conditional alpha-beta model (which also has a relatively low SIC value of 1.359). This is also a three-factor Fama-French model, where all three *beta* factor loadings are time varying depending on a signal provided by the market dividend yield (Z3).

The Fama-French 3 factor model has been selected as the 'best model' for all three categories: unconditional, conditional beta and conditional alpha-beta model. Figures 1, 2 and 3 respectively, show histograms of the cross-section distribution of the actual alphas estimated from these three models, applied to all funds. There is a wide spread of alpha estimates across all three categories with a reasonable number of funds in each of the tails of the distribution

[Figures 1, 2 and 3 here]

5. Empirical Results: Bootstrap Analysis

In this section we present the main findings from the application of the baseline bootstrap procedure. As discussed previously, we impose a minimum requirement of 60 observations for a fund to be included in the analysis. This leaves a sample of 724 funds, of which 307 are non-survivor funds (i.e. have ceased to exist at some point before the end of the sample period), while 417 are survivor funds. Furthermore, by examining the sensitivity of our bootstrap results to the inclusion of funds with different minimum fund histories, we are able to evaluate whether a possible 'look-ahead' bias is induced by the minimum 60 observation restriction.

In Table 1, we reported that around 70% of mutual funds reject normality in their regression residuals. As this finding partly motivates the use of the nonparametric bootstrap, we provide further information on this aspect. Figures 4a and 5a show the distribution of the residuals for selected funds in the upper (figure 4a) and lower (figure 5a) sections of the (ex-post) performance distribution.

'Mutual Fund Performance: Skill or Luck ?'

Residuals from funds in the extreme tails exhibit higher variance and a greater degree of non-normality than residuals from funds closer to the centre of the performance distribution. For funds in the upper tail, it is this high variance and in particular the existence of large positive residuals, that causes these 'best funds' to populate the very top end of the bootstrap distributions and generates wide dispersion and non-normality among the very 'top performing' funds. This is evident in figure 4b which shows the bootstrap histogram of $t_{\hat{\alpha}_i}$ at selected points of the performance distribution.

Figures 5a and 5b present an almost mirror image for the lower end of the performance distribution. This vividly illustrates that although funds in the center of the performance distribution may exhibit near normal idiosyncratic risks, those in the tails do not and it is the latter in which investors are particularly interested.

[Figure 4a and 4b here]

[Figure 5a and 5b here]

Table 2 shows bootstrap results for the full set of mutual funds (i.e. including all investment objectives) for the unconditional (Panel A), conditional-beta (Panel B) and conditional alpha-beta (Panel C) models all of which use the Fama-French (FF) three-factor model. The first row in each panel shows the fund's alpha, (% per month) ranked from lowest to highest (left to right) and its associated t-statistic (the second row, 't-alpha'). Row 3 ("t-stat") presents the 'conventional' *t-statistics* of alpha ranked from lowest to highest. Rows 1-3 are obtained using the 'real data'. Row 4 ("p-stat") reports the bootstrap p-values of the t-statistic in row 3, based on the 'luck distribution' for $t_{\hat{\alpha}}$ under the null of no outperformance.

[Table 2 here]

Table 2, Panel A (final column) reveals that the best fund ranked by alpha from the *unconditional* model achieved an abnormal performance of 0.745% p.m. (8.94% p.a.) with an associated 'conventional' t-statistic of 2.546. However, ranking funds by $t_{\hat{\alpha}}$ rather than $\hat{\alpha}$ gives $t_{\hat{\alpha}} = 4.023$ (row 3) for the best fund, with a bootstrap p-value of 0.056⁵. The latter indicates that from among the 1,000 bootstrap simulations across funds, under the null hypothesis of zero abnormal performance, 5.6% of the highest bootstrap t-statistics were greater than 4.023. This

⁵ Of course, the 'best fund' ranked by $t_{\hat{\alpha}}$ will in general be a different fund to that ranked by $\hat{\alpha}$. However, qualitatively similar results are obtained when alpha rather than t-alpha is used as the performance criterion.

'Mutual Fund Performance: Skill or Luck ?'

can be seen in the histogram in the upper left of figure 4b, where the vertical line shows the actual $t_{\hat{\alpha}} = 4.023$, relative to the 'luck distribution. Thus using a 5% upper tail cut off point we marginally fail to reject the hypothesis that the best fund's $t_{\hat{\alpha}} = 4.023$ may be explained by luck alone.

However, there is strong evidence in table 2 ('p-tstat') to indicate that somewhat lower ranked funds (e.g. 10th, 12th, 15th and 20th) are skilled in selecting stocks with only a 2% chance or less that their positive $t_{\hat{\alpha}}$'s are due to luck. There appears to be genuine skill and positive abnormal performance for these funds. Results (not reported) show that all funds ranked between 10th and 21st inclusive beat the bootstrap estimate of luck at a 95% confidence level. Of the top 20 ranked funds, 16 achieve genuine outperformance at 90% confidence while 11 funds outperform at 95% confidence. However, as one moves slightly closer to the centre of the performance distribution (i.e. at or below the top 3% level) there is no evidence of stock picking ability – any positive $t_{\hat{\alpha}}$'s are due to luck rather than skill.

In the left tail of the distribution, (i.e. the left side of Panel A, table 2), the worst ranked fund by alpha yields a negative return of -0.901% p.m. (-10.8% p.a.) with a conventional t-statistic of -2.532. The lowest ranked fund ordered by $t_{\hat{\alpha}}$ yields a t-statistic of -7.414 with a bootstrap p-value of 0.000. Hence for the ex-post 'worst fund' with $t_{\hat{\alpha}} = -7.414$, there is a near zero probability that this is due to bad luck rather than 'bad skill'. This fund has produced 'truly' inferior performance. This can be seen in the upper left panel of figure 5b, where the vertical line indicates an actual $t_{\hat{\alpha}} = -7.414$, which is to the left of the 'luck distribution'. It is clear from Panel A, table 2 (and figure 5b), that all funds in the left tail (up to the 'min 40%' point) have genuinely 'poor skill'.

An alternative interpretation of the bootstrap results is to see how many funds one might expect to achieve a given level of alpha performance by random chance alone and compare this with the number of funds which actually did achieve this level of alpha in the 'real world'. For example, based on the unconditional (FF) model we would expect 6 funds to achieve $\hat{\alpha} \geq 0.5\%$ p.m. (6% p.a.) based on random chance alone, whereas 13 funds exhibit this level of performance (or higher). However, $\hat{\alpha} \geq 0.1\%$ p.m. (1.2% p.a.) is expected to be achieved by 171 funds solely based on chance, while in fact only 133 funds are observed to have reached this level of performance. This interpretation is, of course, consistent with the discussion of p-values

'Mutual Fund Performance: Skill or Luck ?'

above. There is greater evidence of genuine outperformance *just within* the extreme right tail, than nearer the centre of the performance distribution.

[Figure 6 here]

Figure 6 reinforces the above point by showing Kernel density estimates of the distributions of $t_{\hat{\alpha}}$ in the 'real data' and the bootstrap distribution for $t_{\hat{\alpha}}$ - under the null of zero outperformance (i.e. the 'luck distribution'). It shows that the left tails of the distribution of $t_{\hat{\alpha}}$ using the 'real data' (dashed line), lies largely to the left of the bootstrap distribution (continuous line). Such poor performing funds cannot attribute their performance to bad luck but have 'poor skill'. In contrast, the right tail of the distribution of $t_{\hat{\alpha}}$ for the 'real data' lies outside the 'luck distribution' for only a few funds. This means there are some, but not many, genuine 'outperformers'.

Panel B of Table 2 reports findings from the conditional-beta FF model. Inferences from the bootstrap (rows 't-stat', 'p-stat'), for both the left and right tail of the performance distribution are very similar to those for the unconditional FF model in Panel A. Our results are qualitatively consistent with Kosowski et al (2004) who find strong evidence of stock picking ability among top performing US funds (but somewhat more pervasive than in our UK study) and genuine poor performance for the funds in much of the left tail of the performance distribution.

Inferences for the conditional 'alpha-beta' FF model (Panel C, Table 2) are also consistent with the previous two classes of model. Some genuine stock picking ability is found among some funds at the top end of the performance distribution but skill is distinctly lacking among funds at the lower end where bad luck is not a defence for bad performance.

Above we applied the bootstrap to our 3 'best models'. However, recall from Table 1 (Panel B and C) that the set of conditioning information variables were shown to be only weakly statistically determined (on average across funds) in all of the alternative conditional models shown (and in many further specifications not shown) - and these variables are also statistically insignificant for more than 90% of the funds. Therefore, there is little evidence that conditional models offer additional explanatory power or are likely candidates for the 'true' equilibrium model of returns. We are inclined to place greater weight on results from the unconditional FF model of panel A and our variants (described below) use this 'baseline model'.

Performance and Investment Styles

'Mutual Fund Performance: Skill or Luck ?'

Having found some 'good skill' and lots of 'bad skill' when analyzing all UK mutual funds together, the question now arises whether those skillful and not so skillful funds are equally distributed across different fund classifications or whether they are concentrated in particular investment styles. From the US mutual fund performance literature, there is some evidence that funds with a 'growth' investment style tend to be among the top performing funds (see Chen, Jegadeesh and Wermers 2000).

In our data set 724 funds have a minimum of 60 monthly observations of which 192 are income funds (27%), 82 are growth funds (11%), 326 are general equity (income and growth) funds (45%) and 124 are small stock funds (17%). To further address the 'style question' we apply the bootstrap procedure separately for each style. Here, one is examining a more homogenous risk group and this helps to control for possible unknown cross-sectional risk characteristics, which may be inadequately specified by the equilibrium model of returns.

[Table 3 here]

Table 3, Panels A, B, C and D re-estimate the performance statistics of table 2, for the four investment styles. Looking at all four Panels in Table 3 ('t-stat', 'p-stat') it is clear that genuine 'bad skill' in the left tail is common across all styles but 'good skill' in the right tail differs across styles. Many high ranking equity income funds and equity growth funds (Panel A and Panel B, respectively) in the right tail achieve positive levels of performance which cannot be accounted for by luck. In particular, (in results not shown) the equity income funds ranked from the 4th highest to the 26th highest beat the bootstrap estimate of luck (at a 5% significance level), while the performance of higher ranked funds could have been achieved by luck alone. Similarly, among equity growth funds, while one cannot reject that performance at the extreme top end (i.e. 1st and 2nd ranked funds) could be attributed to good luck, the performance of slightly lower ranked funds is due to genuine skill.

In contrast to the above, for general equity and small stock funds (Table 3, Panel C and Panel D, 't-stat', 'p-stat'), there is virtually no evidence of genuine stock picking skill in the right tail of the performance distribution.

Our findings for the UK of 'skill' among some income funds and growth funds and poor performance in the left tail for the income, growth and general equity styles is broadly consistent with Kosowski et al (2004) for US mutual funds, although the latter study fails to reject the hypothesis that many poor performing US aggressive growth stock funds are merely 'unlucky'.

[Figure 7 here]

Figure 7 shows Kernel density estimates for the four investment styles. We see that the left tails of the actual ex-post t-statistics lie largely to the left of the 'luck distributions', indicating that poor performance is unlikely to be due to bad luck in each of the four investment styles. For equity income funds and equity growth funds, the extreme right tail of the distribution of actual t-statistics lies outside that of the 'luck distribution', indicating the presence of some funds with 'good skill' rather than good luck. But for general equity and small stock funds, actual ex-post performance does not exceed random sampling variation⁶.

Performance and Fund Location

All mutual funds in this study invest only in UK equity but funds are operated from both onshore UK and offshore locations such as Dublin, Luxembourg, Channel Islands and some other European locations. Differential performance may arise due to possible information asymmetries between offshore versus onshore operations or simply differential skill given identical information.

[Table 4 here]

[Figure 8 here]

The bootstrap results in table 4 ('t-stat', 'p-stat') are clear cut. Almost all of the top 20 onshore funds possess genuine skill (panel A), whereas any positive abnormal performance by offshore funds (Panel B) may be attributed to luck. For the lower end of the performance distribution, both onshore and offshore funds demonstrate 'poor skill'. These results are clearly demonstrated in the Kernel density estimates of the actual and bootstrapped t-statistics which are shown in figure 8 for onshore and offshore funds, respectively.

Performance and Survival

As previously outlined, to mitigate survivor bias we include 307 nonsurvivor funds among the full set of 724 in the bootstrap analysis. Nonsurvivor funds are well represented among funds whose performance is superior to chance at 90% and even 95% confidence. For example, among the top 20 ranked funds by the unconditional FF model, 11 funds beat luck at 95%

⁶ It should be noted that Kernel density plots need not necessarily lead to the same conclusion as the bootstrap analysis. This is because the Kernels compare the frequency of a *given level of performance* from among the actual funds, against the frequency of this same level of performance in the *entire* bootstrap matrix. The bootstrap p-value is a more sophisticated measure and compares the actual performance measure $t_{\hat{\alpha}}$ against the bootstrap distribution of performance $t_{\tilde{\alpha}}$, *at the same point* in the cross-sectional performance distribution.

'Mutual Fund Performance: Skill or Luck ?'

confidence, 5 of which are nonsurvivors while 16 funds beat luck at 90% confidence, 9 of which are nonsurvivors. This class of funds are not an inferior class of funds.

The restriction of a minimum of 60 observations for fund inclusion in the analysis, while statistically advantageous, may induce a 'look-ahead bias' (i.e. requiring funds to survive a minimum period of time – Carhart et al 2002). To address this issue we repeat the bootstrap procedure for alternative minimum fund histories of 18, 36, 60 and 120 observations respectively.

[Table 5 here]

Using the FF 3 factor model and all funds, the 'alpha' and 't-alpha' (Table 5) reveal that imposing a minimum observation restriction does not bias the performance findings *upwards*. Kosowski et al (2004) report similar findings. On the contrary, we find some evidence using the ex-post 'real data' that positive abnormal performance (as measured by both alpha and the t-statistic of alpha) is generally stronger, as the minimum fund history requirement is reduced from 120 to 60 to 36 to 18. Also the bootstrap p-values of $t_{\hat{\alpha}}$ ('p-stat') suggest that genuine stock picking ability becomes somewhat more prevalent as minimum fund histories become shorter. From the four panels in Table 5, it is apparent that 'bad skill' is independent of the 'life-span' of the fund (i.e. minimum sample size used)⁷.

Why any fund, particularly a long-lived fund, which truly underperforms would be permitted to survive in a competitive market is puzzling. Kosowski et al (2004) also find strong evidence of genuine inferior performance and argue that this may be because performance measurement is a difficult task requiring, for precision, a long fund life-span. Hendricks et al (1993) suggest that sustained inferior performing funds are those without skill which "churn" their portfolios too much and thus incur relatively high expenses which lowers their performance.

A possible explanation for the unexpected strong performance of nonsurviving funds is that a large number of this class of funds were not forced to close down due to bad performance but in fact were merged or taken over, possibly because of their strong performance and consequent attractiveness. It may be that shorter-lived funds are initially set up to exploit 'new' perceived investment styles and these successful funds are then taken over (possibly by larger

⁷ Looking at the number of observations (final row, table 5) for some of the poorest performing funds we see that they have survived for the full sample period from April 1975 – December 2002 (333 monthly observations), whereas some of the best performing funds are relatively short-lived. We find that when the minimum number of observation is set to 18 observations, 5 of the 10 lowest ranked funds have more than 300 observations and 7 of these 10 funds have more than 200 observations. In contrast, of the top 10 ranked funds only 1 fund has more than 200 observations.

'Mutual Fund Performance: Skill or Luck ?'

established funds). Indeed, Blake and Timmermann (1998), point out that 89% of the mutual funds reported as nonsurvivor funds were merged with other funds while only 11% were closed down over the period. A large number of such 'mergers' may be due to good rather than bad performance.

As alpha, $\hat{\alpha}$, and the t-statistic of alpha, $t_{\hat{\alpha}}$, in effect measure average performance over the life of a fund, these measures may disguise a sharp deterioration in performance shortly before the closure of a fund. However, in results not shown but available on request, we compare the performance of survivor funds with nonsurvivor funds in each month over the final two years of nonsurvivor funds' 'lives'. We find no evidence that funds commonly labeled 'nonsurvivors' underperform survivor funds.

Extensions of the Bootstrap

The 'baseline' bootstrap procedure described in section 3 can be modified to incorporate further characteristics of the return data, for example, serial correlation, heteroscedasticity or cross-sectional (across funds) correlations among the idiosyncratic component of returns. Where fund regression residuals indicate that such features are present, refinements to the bootstrap procedure help to retain this information in the construction of the bootstrap 'luck' distributions. This is important in order to mimic the underlying 'true' return generating process as closely as possible. In an appendix to this study, available on request from the authors, we describe these bootstrap extensions in more detail and present results from their implementation. However, we find that inferences regarding skill versus luck in performance reported here are unaltered.

6. Conclusion

Using a comprehensive data set for the UK, April 1975 – December 2002, with over 1500 mutual funds, we use a bootstrap methodology to distinguish between 'skill' and 'luck'. Our study points to the existence of genuine stock picking ability among a *small number* of top performing UK equity mutual funds (i.e. performance which is not solely due to good luck). This is consistent with Kosowski et al (2004) although in the latter case skill is found to be more prevalent, particularly in the extreme tails.

Controlling for different investment objectives among funds, it is found that some of the top ranked equity income funds and equity growth funds show genuine stock picking skills whereas such ability is not found among small stock funds and general equity funds. We also

'Mutual Fund Performance: Skill or Luck ?'

find that positive performance amongst onshore funds is due to genuine skill, whereas for offshore funds, positive performance is attributable to luck.

At the negative end of the performance scale, our *whole analysis* strongly rejects the hypothesis that poor performing funds are merely unlucky. These funds demonstrate 'bad skill', which is consistent with results for US funds (Kosowski et al 2004).

Our results are robust with respect to alternative equilibrium models, different bootstrap resampling methods and allowing for the correlation of idiosyncratic shocks both within and across funds.

The basic bootstrap is also applied to funds using alternative minimum fund histories, ranging from minimum 18 to 120 observations. There is evidence of a 'reverse look-ahead bias' in that inclusion of shorter lived funds increases the prevalence of genuine abnormal positive performance – we suggest a number of possible explanations for this.

For the active fund management industry as a whole, our findings are something of a curate's egg. It is possible for sophisticated investors to pick funds whose risk adjusted return is due to skill and who could therefore justify higher management fees. However, for the vast majority of funds with positive abnormal performance, we find this is purely due to good luck. We also show that the 'genuine' top performers are not necessarily those with an ex-post ranking right at the 'top'. This makes it extremely difficult for the 'average investor' to pinpoint active funds which demonstrates genuine skill, based on their track records.

The continued existence over long time periods, of a large number of funds which have a truly inferior performance (which cannot be attributed to bad luck), together with many funds who have a positive performance due to good luck (rather than good skill), indicates that many investors either cannot correctly evaluate fund performance or, find it 'costly' to switch between funds.

The above conclusions suggest that at the present time many UK equity investors would be far better investing in index/tracker funds, with their lower transactions costs. On the policy side the UK government seems likely to want to encourage long term saving via mutual (and pension) funds (Turner 2004). So perhaps it is about time the UK government increased its 'health warning' on investing in equity funds from, 'The value of your investments can go down as well as up' to, 'Active fund managers can damage your financial health'. Of course the above is

'Mutual Fund Performance: Skill or Luck ?'

predicated on the risk models examined in this study but we have explored many variants, using data on many funds (both survivors and non-survivors) and a long fund history.

References

- Allen , D.E. and Tan, M.L. (1999) "A Test of the Persistence in the Performance of UK Managed Funds", *Journal of Business Finance and Accounting*, 25(5)&(6), 559-593.
- Ang, A. and Bekaert, G. (2004) 'Stock Return Predictability – Is it There?' *Working Paper, Columbia University, New York*
- Baks, K.P., Metrick, A. and Wachter, J. (2001) 'Should Investors Avoid All Actively Managed Funds? A Study in Bayesian Performance Evaluation', *Journal of Finance*, 56(1), 45-86.
- Ball, R. and Kothari, S.P. (1989) 'Nonstationary Expected Returns: Implications for Tests of Market Efficiency and Serial Correlations in Returns', *Journal of Financial Economics*, 25, 51-74.
- Barberis, N. (2000) 'Investing for the Long Run When Returns are Predictable', *Journal of Finance*, 55, 225-264.
- Barberis, N. and Thaler, R. (2003) 'A Survey of Behavioral Finance', in *Handbook of the Economics of Finance*, edited by G.M. Constantinidis, M. Harris and R. Stulz, Elsevier Science B.V.
- Benartzi, S. and Thaler, R.H. (2001) 'Naïve Diversification Strategies in Defined Contribution Savings Plans', *American Economic Review*, 91(1), 79-88.
- Berk, J.B. and Green, R.C. (2004) 'Mutual Fund Flows and Performance in Rational Markets', *Journal of Political Economy*, 112(6), 1269-95.
- Bildersee, J.S. (1975) 'The Association Between a Market Determined Measure of Risk and Other Measures of Risk', *Accounting Review*, 50, 81-98.
- Blake, C.A. and Morey, M (2000) 'Morningstar Ratings and Mutual Fund Performance', *Journal of Financial and Quantitative Analysis*, 35(3), 451-483.
- Blake, D. and Timmermann, A. (1998) 'Mutual Fund Performance : Evidence from the UK', *European Finance Review*, 2, 57-77.
- Blake, D. and Timmermann, A. (2003) 'Performance Persistence in Mutual Funds : An Independent Assessment of the Studies Prepared by Charles River Associates for the Investment Management Association', *Financial Services Authority (also available at www.pensions-institute.org/reports)*
- Blake, D., Lehmann, B. and Timmermann, A. (1999) "Performance Measurement Using Multi-Asset Portfolio Data: A Study of UK Pension Funds 1986-94, Pensions Institute, London.
- Bodie, Z., Merton, R.C. and Samuelson, W. (1992) 'Labour Supply Flexibility and Portfolio Choice in a Life-Cycle Model', *Journal of Economic Dynamics and Control*, 15, 725 – 31.
- Bollen, N.P.B. and Busse, J.A. (2005) 'Short-Term Persistence in Mutual Fund

'Mutual Fund Performance: Skill or Luck ?'

- Performance', *Review of Financial Studies*, forthcoming.
- Brennan, M.J., Schwartz, E.S. and Lagnado, R. (1997) 'Strategic Asset Allocation', *Journal of Economic Dynamics and Control*, 2, 1377-1403
- Brown, S. and Goetzmann, W (1995) "Performance Persistence", *Journal of Finance*, 50(2), 679-698.
- Campbell, J.Y. and Viceira, L.M. (1999) 'Consumption and Portfolio Decisions when Expected Returns are Time Varying', *Quarterly Journal of Economics*, 114(2), 433-496
- Campbell, J.Y., Chan, Y.L. and Viceira, L.M. (2003) 'A Multivariate Model of Strategic Asset Allocation', *Journal of Financial Economics*, 67(1), 41-80
- Carhart, M. (1997) 'On Persistence in Mutual Fund Performance', *Journal of Finance*, 52(1), 57-82
- Carhart, M., Carpenter, J., Lynch, A. and Musto, D. (2002) 'Mutual Fund Survivorship', *The Review of Financial Studies*, 15(5), 1439-1463.
- Chan, K.C. (1988) 'On the Contrarian Investment Strategy', *Journal of Business*, 61(1), 147-164.
- Chan, L. and Lakonishok, J. (2004) 'Value and Growth Investing: Review and Update', *Financial Analysts Journal*, 60, 71-86.
- Chan, L.K.C, Karceski, J. and Lakonishok, J. (2000) 'New Paradigm or Same Old Hype in Equity Investing?', *Financial Analysis Journal*, 56, 23-36
- Chan, L.K.C., Jegadeesh, N. and Lakonishok, J, (1996) 'Momentum Strategies', *Journal of Finance*, 51(5), 1681-1713.
- Charles River Associates (2002) 'Performance Persistence in UK Equity Funds – An Empirical Analysis', prepared by T. Giles, T. Wilsdon and T. Worboys, October.
- Chen, H.L., Jegadeesh, N. and Wermers, R. (2000) 'The Value of Active Mutual Fund Management : An Examination of the Stockholdings and Trades of Fund Managers', *Journal of Financial and Quantitative Analysis*, Vol. 35, No. 3, pp. 343-368.
- Chevalier, J. and Ellison, G. (1999) 'Are Some Mutual Fund Managers Better than Others? Cross-Sectional Patterns in Behavior and Performance', *Journal of Finance*, 54(3), 875-899.
- Christopherson, J. Ferson, and W. Glassman, D. (1998) 'Conditioning Manager Alphas on Economic Information: Another Look at the Persistence of Performance', *Review of Financial Studies*, Vol. 11, No. 1, pp. 111-142
- Daniel, K. M., Grinblatt, M., Titman, S. and Wermers, R (1997) "Measuring Mutual Fund Performance With Characteristic Based Benchmarks", *Journal of Finance*, 52, 1035-1058.
- Efron, B. and Tibshirani R.J. (1993). *An Introduction to the Bootstrap*, Monographs on Statistics and Applied Probability, Chapman and Hall, New York.

'Mutual Fund Performance: Skill or Luck ?'

- Fama, E.F. and French, K.R. (1993) 'Common Risk Factors in the Returns on Stocks and Bonds', *Journal of Financial Economics*, 33, 3-56.
- Ferson, W. and Schadt, R. (1996) 'Measuring Fund Strategy and Performance in Changing Economic Conditions', *Journal of Finance*, 51, 425-62.
- Fletcher, J. (1995) 'An Examination of the Selectivity and Market Timing Performance of UK Unit Trusts', *Journal of Business Finance and Accounting*, 22(1), 143-156.
- Fletcher, J. and Forbes, D. (2002) 'An Exploration of the Persistence of UK Unit Trusts Performance', *Journal of Empirical Finance*, 9(5), 475-493.
- Foster, G. (1986) *Financial Statement Analysis*, Prentice Hall, Englewood Cliffs.
- Goetzmann, W., Ivkovic, Z. and Rouwenhorst, G. (2001) "Day Trading International Mutual Funds: Evidence and Policy Solutions" *Journal of Financial and Quantitative Analysis*, 36(3), 287-310.
- Grinblatt, M., Titman, S. and Wermers, R (1995) "Momentum Investment Strategies, Portfolio Performance and Herding: A Study of Mutual Fund Behavior, *American Economic Review*, 85, 1088-1105.
- Hall, P. (1992) *The Bootstrap and Edgeworth Expansion*, Springer Verlag
- Hendricks, D., Patel, J. and Zeckhauser, R. (1993) 'Hot Hands in Mutual Funds: Short Run Persistence of Performance, 1974-88', *Journal of Finance*, 48, 93-130.
- Hochman, S. (1983) 'The Beta Coefficient: An Instrumental Variable Approach', *Research in Finance*, 4, 392-407.
- Jegadeesh, N. and Titman, S. (1993) 'Returns to Buying Winners and Selling Losers : Implications for Stock Market Efficiency', *Journal of Finance*, 48(1), 56-91.
- Jegadeesh, N. and Titman, S. (2001) 'Profitability of Momentum Strategies : Evaluation of Alternative Explanations', *Journal of Finance*, 56(2), 699-720.
- Jensen, M. (1968) 'The Performance of Mutual Funds in the Period 1945-1964', *Journal of Finance*, 23, 389-416.
- Jiang, W. (2003) 'A Non Parametric Test of Market Timing', *Journal of Empirical Finance*, 10, 399-425.
- Kosowski, R., Timmermann, A., White, H. and Wermers, R. (2004) 'Can Mutual Fund "Stars" Really Pick Stocks? New Evidence from a Bootstrap Analysis', *Discussion Paper*, INSEAD, April.
- LaPorta, R., Lakonishok, J., Shleifer, A. and Vishny, R. (1997) 'Good News for Value Stocks: Further Evidence on Market Efficiency' *Journal of Finance*, 52(2), 859-874
- Lakonishok, J.A., Shleifer, A. and Vishny, R.W. (1992) 'The Structure and Performance of the Money Management Industry', *Brookings Papers on Economic Activity*, 339-391.
- Leger, L. (1997) 'UK Investment Trusts : Performance, Timing and Selectivity', *Applied Economics Letters*, 4, 207-210.

'Mutual Fund Performance: Skill or Luck ?'

- Mamaysky, H., Spiegel, M. and Zhang, H (2004) 'Improved Forecasting of Mutual Fund Alphas and Betas', Yale School of Management, ICF Working Paper 04-23.
- Mandelker, G.N. and Rhee, S.G. (1984) 'The Impact of the Degrees of Operating and Financial Leverage on Systematic Risk of Common Stock', *Journal of Financial and Quantitative Analysis*, 19, 45-57.
- Merton, R. and Henriksson, R. (1981) 'On Market Timing and Investment Performance : Statistical Procedures for Evaluating Forecasting Skills', *Journal of Business*, 54, 513-533
- Myners, P. (2001) 'Institutional Investment in the United Kingdom: A Review', *Report prepared for the Chancellor of the Exchequer*, H.M Treasury, London.
- Newey, W. and West, K. (1987) 'A Simple, Positive Semi-Definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix', *Econometrica*, 55, 703-708.
- OECD (2003) 'Monitoring the Future Social Implication of Today's Pension Policies', *OECD*, Paris, unpublished.
- Pastor, L. and Stamburgh, R. (2002) "Mutual Fund Performance and Seemingly Unrelated Assets" *Journal of Financial Economics*, 63(3), 315-350.
- Pesaran, M.H. and Timmermann, A. (1994) 'Forecasting Stock Returns: An Examination of Stock Market Trading in the Presence of Transaction Costs', *Journal of Forecasting*, 13(4), 335-367
- Pesaran, M.H. and Timmermann, A. (1995) 'Predictability of Stock Returns : Robustness and Economic Significance', *Journal of Finance*, 50(4), 1201-1228.
- Pesaran, M.H. and Timmermann, A. (2000) 'A Recursive Modelling Approach to Predicting UK Stock Returns' *Economic Journal*, 110(460), 159-191.
- Politis, D.N. and Romano, J.P. (1994) 'The Stationary Bootstrap', *Journal of the American Statistical Association*, 89, 1303-1313.
- Presidential Commission on Social Security Reform (2001) "Strengthening Social Security and Creating Personal Wealth for All Americans", *Report of the President's Commission*, December 2001, Washington DC (available at www.csss.gov).
- Quigley, G. and Siquefield, R.A. (2000) 'Performance of UK Equity Unit Trusts', *Journal of Asset Management*, 1(1), 72-92
- State of the Union Address (2005), Chamber of the U.S. House of Representatives, The United States Capitol, Washington, D.C. (available at www.whitehouse.gov).
- Teo, M. and Woo, Sung-Jun (2001) "Persistence in Style-Adjusted Mutual Fund Returns", Manuscript, Harvard University.
- Thomas, A. and Tonks, I (2001) 'Equity Performance of Segregated Pension Funds in the UK', *Journal of Asset Management*, 1(4), 321-343.
- Tonks, I. (2004) "Performance Persistence of Pension Fund Managers", University of Exeter, Centre for Finance and Investment, Working Paper.

'Mutual Fund Performance: Skill or Luck ?'

- Treynor, J. and Mazuy, K. (1966) 'Can Mutual Funds Outguess the Market', *Harvard Business Review*, 44, 66-86.
- Turner, A. (2004) 'Pensions : Challenges and Choices : The First Report of the Pensions Commission', The Pensions Commission, *The Stationary Office*, London
- Viceira, L.M. (2001) 'Optimal Portfolio Choice for Long-horizon Investors with Nontradable Labor Income', *Journal of Finance*, 56(2), 433-470.
- Wermers, R. (2000) 'Mutual Fund Performance : An Empirical Decomposition into Stock Picking Talent, Style, Transactions Costs, and Expenses', *Journal of Finance*, 55(4), 1655-1703.
- Wermers, R.(2003) 'Is Money Really "Smart"? New Evidence on the Relation Between Mutual Fund Flows and Performance Persistence', Department of Finance, University of Maryland.
- Xia, T. (2001) 'Learning About Predictability : The Effects of Parameter Uncertainty on Dynamic Asset Allocation', *Journal of Finance*, 56(1), 205-246.

Table 1. Model Selection: Cross-Sectional Results of Model Estimations.

Table 1 shows results from the estimation of the performance models described in Section II using all mutual funds. Panel A reports unconditional models, Panel B conditional-beta models and Panel C conditional alpha-beta models. The t-statistics are based on Newey-West heteroscedasticity and autocorrelation adjusted standard errors. (t-statistics shown are cross-sectional averages of the absolute value of funds t-statistics). Also shown are statistics on the percentage of funds which (i) reject normality in the residuals (Jarque-Bera test) and (ii) reject the null hypothesis of no serial correlation in residuals at lags 1 to 6 (LM test) – both at 5% significance level. Also shown is the Schwartz Information Criterion (SIC). The table also shows alpha and its t-statistic, for an equal weighted portfolio of all mutual funds. All figures shown are cross-sectional averages.

Panel A: Unconditional Factor Models							
			1	2	3	4	5
del			CAPM	FF	Carhart	TM	MH
Regression Coefficients							
Average Alpha (percent per month)			-0.29	-0.069	-0.072	0.086	0.228
t-statistic			1.022	1.251	1.283	1.135	1.212
Standard Deviation of Alpha			0.223	0.219	0.222	0.300	0.418
Unconditional Betas							
(t-statistics in parentheses)	r_{mt}		0.889 (24.318)	0.918 (27.485)	0.921 (27.93)	0.878 (24.08)	0.956 (16.76)
	SMB			0.286 (5.366)	0.286 (5.437)		
	HML			-0.008 (1.476)	-0.003 (1.326)		
	PR1YR				0.010 (1.240)		
Market Timing Measures							
Treyner – Mazuy						-0.005 (1.527)	
Merton-Henriksson							-0.149 (1.272)
Model Selection Criteria							
Adjusted R-square			0.743	0.821	0.824	0.747	0.746
Schwartz Information Criterion (SIC)			1.521	1.299	1.309	1.534	1.537
Rejection of Normality (% of funds)			71	72	70	71	71
LM(1) statistic (% of Funds)			45	42	42	44	44
LM(6) statistic (% of Funds)			38	38	39	40	38
Equal Weighted Portfolio							
Alpha			0.027	-0.141	-0.146	0.174	0.370
t-statistic			0.380	-3.110	-3.183	2.291	3.499

'Mutual Fund Performance: Skill or Luck ?'

Panel B: Conditional Beta Models							
		6	7	8	9	10	11
Model		FS	FF(1)	FF(2)	Carhart	TM	FF(3)
Regression Coefficients							
Average Alpha (percent per month)		0.015	-0.042	-0.042	-0.045	0.106	-0.045
t-statistic		1.053	1.210	1.204	1.234	1.147	1.215
Standard Deviation of Alpha		0.251	0.226	0.221	0.224	0.305	0.216
Unconditional Betas							
(t-statistics in parentheses)	r_{mt}	0.822 (18.079)	0.889 (19.912)	0.884 (24.881)	0.883 (20.041)	0.819 (17.901)	0.876 (25.084)
	SMB		0.284 (5.389)	0.283 (5.315)	0.284 (5.473)		0.268 (4.725)
	HML		-0.008 (1.467)	-0.007 (1.476)	-0.004 (1.363)		0.026 (1.321)
	PR1YR				0.009 (1.194)		
Market Timing Measures							
Treynor – Mazuy	r_{mt}^2					-0.005 (1.165)	
Conditioning Variables, Z_{t-1}							
$Z1_{t-1} * r_{mt}$, Z1: One month rate		0.031 (1.232)	0.027 (1.336)		-0.024 (1.305)	-0.093 (1.143)	
$Z2_{t-1} * r_{mt}$, Z2: Term Spread		-0.009 (1.157)	-0.078 (1.281)		-0.126 (1.239)	-0.139 (1.163)	
$Z3_{t-1} * r_{mt}$, Z3: Dividend Yield		-0.074 (1.285)	-0.044 (1.356)	-0.041 (1.420)	-0.038 (1.236)	-0.037 (1.031)	-0.045 (1.451)
$Z3_{t-1} * SMB_t$							0.009 (1.546)
$Z3_{t-1} * HML_t$							0.032 (1.018)
Model Selection Criteria							
Adjusted R-square		0.751	0.828	0.824	0.831	0.753	0.828
Schwartz Information Criterion (SIC)		1.568	1.341	1.309	1.353	1.587	1.338
Rejection of Normality (% of funds)		71	71	71	69	70	71
LM(1) statistic (% of Funds)		44	43	41	43	43	44
LM(6) statistic (% of Funds)		40	42	39	44	42	41
Equal Weighted Portfolio							
Alpha		0.076	-0.107	-0.110	-0.116	0.162	-0.110
t-statistic		1.079	-2.417	-2.465	-2.580	2.177	-2.524

'Mutual Fund Performance: Skill or Luck ?'

Panel C: Conditional Alpha and Beta Models								
		12	13	14	15	16	17	18
Model		FS	FF(1)	FF(2)	FF(3)	FF(4)	FF(5)	Carhart
Regression Coefficients								
Average Alpha (percent per month)		0.030	0.076	-0.056	-0.053	0.057	-0.004	0.068
t-statistic		0.985	1.127	1.185	1.191	1.143	1.161	1.136
Standard Deviation of Alpha		0.637	0.707	0.386	0.386	0.690	0.464	0.654
Unconditional Betas								
(t-statistics in parentheses)	r_{mt}	0.820 (18.020)	0.882 (19.843)	0.886 (24.924)	0.877 (25.125)	0.900 (20.600)	0.921 (25.153)	0.877 (19.978)
	SMB		0.283 (5.385)	0.283 (5.335)	0.268 (4.758)	0.286 (5.395)	0.275 (4.808)	0.283 (5.479)
	HML		-0.010 (1.442)	-0.008 (1.459)	0.024 (1.302)	-0.011 (1.486)	0.030 (1.404)	-0.006 (1.342)
	PR1YR							0.010 (1.917)
Conditioning Variables, Z_{t-1}								
$Z1_{t-1} * r_{mt}$, Z1: One month rate		-0.016 (1.280)	-0.015 (1.374)			-0.120 (1.360)	-0.017 (1.216)	-0.065 (1.345)
$Z2_{t-1} * r_{mt}$, Z2: Term Spread		-0.096 (1.217)	-0.145 (1.314)			-0.248 (1.411)		-0.192 (1.287)
$Z3_{t-1} * r_{mt}$, Z3: Dividend Yield		-0.066 (1.309)	-0.039 (1.374)	-0.040 (1.428)	-0.044 (1.465)			-0.033 (1.247)
$Z3_{t-1} * SMB_t$					0.010 (1.571)			
$Z3_{t-1} * HML_t$					0.030 (1.016)			
$Z1_{t-1} * SMB_t$							0.064 (1.547)	
$Z1_{t-1} * HML_t$							0.125 (1.005)	
$Z1_{t-1}$		1.259 (0.961)	1.167 (0.973)			0.285 (0.968)	0.120 (0.945)	1.158 (0.969)
$Z2_{t-1}$		2.340 (1.110)	1.161 (0.944)			0.304 (0.898)		1.134 (0.932)
$Z3_{t-1}$		-0.317 (1.020)	-0.203 (0.975)	-0.049 (0.965)	-0.042 (1.026)			-0.204 (0.982)
Model Selection Criteria								
Adjusted R-square		0.753	0.830	0.825	0.829	0.826	0.826	0.832
Schwartz Information Criterion (SIC)		1.638	1.413	1.331	1.359	1.378	1.375	1.424
Rejection of Normality (% of funds)		72	70	71	70	72	73	68
LM(1) statistic (% of Funds)		45	49	44	48	46	46	50
LM(6) statistic (% of Funds)		42	46	38	42	45	43	47
Equal Weighted Portfolio								
Alpha		0.071	-0.113	-0.114	-0.117	-0.129	-0.125	-0.122
t-statistic		1.008	-2.570	-2.603	-2.724	-2.933	-2.817	-2.727

Table 2 : Bootstrap Results of UK Mutual Fund Performance

Table 2 shows statistics for the full sample of mutual funds (including all investment objectives) for each of the three types of model selected in section 4. Panel A reports statistics from the unconditional Fama and French FF (three-factor) model, Panel B for the conditional-beta FF model and Panel C for the FF conditional alpha-beta model. The first row in each panel reports the ex-post alpha (% per month) for various points and percentiles of the performance distribution, ranging from worst fund (min) to best fund (max). The second row reports the associated t-statistic for these alphas. Row 3 contains the ex-post t-statistics of alpha ranked from lowest (min) to highest (max). Row 4 reports the bootstrap p-values of the t-statistics in row 3 based on 1,000 bootstrap resamples. Both actual ex-post and bootstrap t-statistics are based on Newey-West heteroscedasticity and autocorrelation adjusted standard errors. Results are restricted to funds with a minimum of 60 observations.

Panel A : Unconditional Model																		
$(R_i - r_f)_t = \alpha_i + \beta_{1i}(R_m - r_f)_t + \beta_{2i}SML_t + \beta_{3i}HML_t + \epsilon_{it}$																		
	min	5.min	min5%	min10%	min40%	max40%	max10%	max5%	max3%	20max	15max	12max	10max	7max	5.max	3.max	2.max	max
Alpha	-0.901	-0.663	-0.427	-0.330	-0.108	-0.027	0.181	0.301	0.411	0.439	0.478	0.507	0.530	0.588	0.598	0.686	0.723	0.745
t-alpha	-2.532	-4.445	-1.563	-2.406	-1.591	-0.235	1.757	1.920	3.389	0.806	2.777	2.776	1.698	4.023	2.556	2.991	1.322	2.546
t-stat	-7.414	-4.724	-3.077	-2.537	-0.958	-0.216	1.242	1.670	2.023	2.196	2.403	2.522	2.544	2.671	2.776	2.991	3.389	4.023
p-tstat	0.000	0.000	0.000	0.000	0.000	1.000	0.824	0.508	0.169	0.016	0.004	0.006	0.022	0.066	0.132	0.159	0.059	0.056
Panel B : Conditional-Beta Model																		
$(R_i - r_f)_t = \alpha_i + \beta_{1i}(R_m - r_f)_t + \beta_{2i}SML_t + \beta_{3i}HML_t + \beta_{4i}[Z3_{t-1}(R_m - r_f)_t] + \epsilon_{it}$																		
	min	5.min	min5%	min10%	min40%	max40%	max10%	max5%	max3%	20max	15max	12max	10max	7max	5.max	3.max	2.max	max
Alpha	-0.662	-0.614	-0.376	-0.315	-0.095	-0.004	0.195	0.336	0.444	0.486	0.557	0.580	0.602	0.639	0.731	0.804	0.812	1.123
t-alpha	-2.566	-6.884	-3.353	-2.985	-0.969	-0.044	1.454	1.234	2.265	2.289	0.767	3.144	1.855	2.799	1.072	2.776	3.701	1.803
t-stat	-6.884	-4.149	-3.001	-2.371	-0.798	-0.027	1.362	1.837	2.161	2.289	2.341	2.556	2.650	2.776	2.799	3.701	4.088	4.368
p-tstat	0.000	0.000	0.000	0.000	0.000	1.000	0.193	0.022	0.009	0.003	0.028	0.003	0.002	0.019	0.111	0.001	0.002	0.022
Panel C : Conditional Alpha-Beta Model																		
$(R_i - r_f)_t = \alpha_{\alpha} + \alpha_{\beta}Z3_{t-1} + \beta_{1i}(R_m - r_f)_t + \beta_{2i}SML_t + \beta_{3i}HML_t + \beta_{4i}[Z3_{t-1}(R_m - r_f)_t] + \epsilon_{it}$																		
	min	5.min	min5%	min10%	min40%	max40%	max10%	max5%	max3%	20max	15max	12max	10max	7max	5.max	3.max	2.max	max
Alpha	-2.238	-1.789	-0.544	-0.369	-0.115	-0.017	0.294	0.525	0.734	0.783	0.852	1.073	1.103	1.316	1.468	1.547	1.972	2.229
t-alpha	-1.710	-1.095	-1.508	-2.177	-1.040	-0.162	2.207	1.742	1.532	2.451	1.646	1.328	3.836	2.264	2.873	1.975	3.667	2.986
t-stat	-7.288	-4.504	-2.877	-2.301	-0.847	-0.116	1.328	1.835	2.380	2.451	2.677	2.796	2.873	3.078	3.194	3.667	3.687	3.836
p-tstat	0.000	0.000	0.000	0.000	0.000	1.000	0.552	0.095	0.000	0.000	0.000	0.000	0.002	0.000	0.007	0.005	0.037	0.193

Table 3. Statistical Significance of Mutual Fund Performance by Investment Objective

Table 3 shows statistics for mutual funds categorized by investment objectives, including : Panel A (income funds), Panel B (growth funds), Panel C (general equity - income and growth) funds and Panel D (small stock funds). All results use the unconditional Fama and French three-factor model. The first row in each panel reports the ex-post alpha (% per month) for various points and percentiles of the performance distribution, ranging from worst fund (min) to best fund (max). The second row reports the associated t-statistic for these alphas. Row 3 contains the ex-post t-statistics of alpha ranked from lowest (min) to highest (max). Row 4 reports the bootstrap p-values of the t-statistics in row 3 based on 1,000 bootstrap resamples. Both actual ex-post and bootstrap t-statistics are based on Newey-West heteroscedasticity and autocorrelation adjusted standard errors. Results are restricted to funds with a minimum of 60 observations.

Unconditional Three Factor Model : $(R_i - r_f)_t = \alpha_i + \beta_{1i}(R_m - r_f)_t + \beta_{2i}SML_t + \beta_{3i}HML_t + \varepsilon_{it}$																
Panel A : Equity Income																
	min	5.min	min5%	min10%	min20%	min40%	max40%	max20%	max10%	max7%	max5%	max3%	5.max	3.max	2.max	max
Alpha	-0.666	-0.315	-0.227	-0.164	-0.106	-0.0131	0.027	0.121	0.208	0.240	0.284	0.365	0.382	0.431	0.459	0.478
t-alpha	-7.414	-3.148	-2.313	-1.178	-0.873	-0.240	0.202	1.225	1.623	2.522	2.673	2.196	2.403	2.544	1.726	2.777
t-stat	-7.414	-3.148	-2.469	-1.648	-0.981	-0.318	0.281	0.954	1.623	1.757	1.927	2.403	2.522	2.673	2.777	3.389
p-tstat	0.000	0.000	0.000	0.008	0.148	0.289	0.413	0.154	0.005	0.045	0.046	0.003	0.004	0.058	0.122	0.106
Panel B : Equity Growth																
	min	5.min	min5%	min10%	min20%	min40%	max40%	max20%	max10%	max7%	max5%	max3%	5.max	3.max	2.max	max
Alpha	-0.393	-0.305	-0.305	-0.247	-0.192	-0.141	-0.057	0.065	0.161	0.301	0.377	0.543	0.377	0.543	0.598	0.745
t-alpha	-3.562	-3.146	-3.146	-4.121	-1.486	-1.511	-0.243	0.548	1.466	1.920	2.529	2.671	2.529	2.671	2.556	2.546
t-stat	-4.121	-3.146	-3.146	-2.236	-1.720	-1.129	-0.427	0.424	1.463	1.920	2.282	2.546	2.282	2.546	2.556	2.671
p-tstat	0.011	0.000	0.000	0.000	0.000	0.000	1.000	0.995	0.175	0.042	0.004	0.014	0.004	0.014	0.102	0.333
Panel C : General Equity																
	min	5.min	min5%	min10%	min20%	min40%	max40%	max20%	max10%	max7%	max5%	max3%	5.max	3.max	2.max	max
Alpha	-0.740	-0.650	-0.413	-0.324	-0.241	-0.130	-0.044	0.055	0.135	0.199	0.280	0.439	0.514	0.593	0.606	0.723
t-alpha	-3.043	-2.970	-3.751	-1.937	-2.060	-0.951	-0.624	0.578	0.889	0.778	1.529	0.806	2.127	1.450	2.292	1.322
t-stat	-5.166	-4.118	-3.160	-2.706	-2.056	-1.190	-0.418	0.426	1.053	1.326	1.427	1.627	2.127	2.500	2.776	4.023
p-tstat	0.019	0.000	0.000	0.000	0.000	0.000	1.000	1.000	0.997	0.972	0.995	0.985	0.746	0.452	0.325	0.025
Panel D : Smaller Companies																
	min	5.min	min5%	min10%	min20%	min40%	max40%	max20%	max10%	max7%	max5%	max3%	5.max	3.max	2.max	max
Alpha	-0.901	-0.528	-0.520	-0.469	-0.419	-0.250	-0.105	0.055	0.215	0.314	0.317	0.494	0.466	0.530	0.546	0.686
t-alpha	-2.532	-2.234	-3.077	-2.680	-2.596	-1.135	-0.489	0.228	1.176	0.473	1.610	2.407	1.788	1.698	1.578	2.991
t-stat	-4.445	-3.095	-2.915	-2.739	-2.475	-1.506	-0.470	0.228	0.792	1.242	1.570	1.698	1.610	1.788	2.409	2.991
p-tstat	0.003	0.000	0.000	0.000	0.000	0.000	1.000	1.000	1.000	0.941	0.654	0.857	0.850	0.877	0.338	0.225

Table 4. Statistical Significance of Mutual Fund Performance by Fund Location

Table 4 shows statistics for mutual funds by location. Panel A reports results for funds operated from 'Onshore UK' and Panel B for funds operated 'Offshore'. All results use the unconditional Fama and French three-factor model. The first row in each panel reports the ex-post alpha (% per month) for various points and percentiles of the performance distribution, ranging from worst fund (min) to best fund (max). The second row reports the associated t-statistic for these alphas. Row 3 contains the ex-post t-statistics of alpha ranked from lowest (min) to highest (max). Row 4 reports the bootstrap p-values of the t-statistics in row 3 based on 1,000 bootstrap resamples. Both actual ex-post and bootstrap t-statistics are based on Newey-West heteroscedasticity and autocorrelation adjusted standard errors. Results are restricted to funds with a minimum of 60 observations.

Panel A : Onshore UK funds																	
	min	5.min	min3%	min5%	min10%	min40%	max40%	max10%	max5%	20max	15max	10max	7max	5.max	3.max	2.max	max
Alpha	-0.673	-0.650	-0.469	-0.421	-0.327	-0.100	-0.018	0.199	0.310	0.431	0.470	0.514	0.546	0.598	0.686	0.723	0.754
t-alpha	-2.050	-2.970	-2.680	-2.356	-0.974	-1.647	-0.070	1.617	1.050	2.544	1.077	2.127	1.578	2.556	2.991	1.322	2.546
t-stat	-7.414	-4.724	-3.160	-2.919	-2.464	-0.836	-0.128	1.286	1.726	2.127	2.403	2.544	2.671	2.776	2.991	3.389	4.023
p-tstat	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.576	0.241	0.015	0.001	0.013	0.034	0.083	0.134	0.049	0.036
Panel B : Offshore funds																	
	min	5.min	min3%	min5%	min10%	min40%	max40%	max10%	max5%	20max	15max	10max	7max	5.max	3.max	2.max	max
Alpha	-0.901	-0.523	-0.659	-0.523	-0.413	-0.237	-0.110	0.074	0.228	-0.032	0.006	0.074	0.131	0.228	0.284	0.315	0.593
t-alpha	-2.532	-3.765	-2.284	-3.765	-3.751	-1.209	-1.062	0.489	2.245	-0.119	0.066	0.489	0.896	2.245	1.390	0.406	1.450
t-stat	-4.190	-3.522	-3.765	-3.522	-3.103	-1.707	-0.862	0.406	1.171	-0.125	0.062	0.406	0.654	1.171	1.450	1.451	2.245
p-tstat	0.006	0.000	0.000	0.000	0.000	0.000	1.000	1.000	0.996	1.000	1.000	1.000	1.000	0.996	0.978	0.997	0.751

Table 5 : Performance by Length of Fund Histories

Table 5 reports performance for alternative minimum number of observations. All results use the unconditional FF model and are estimated over all investment objectives. Minimum fund histories vary from 1.5 years to 10 years. In each case fund histories are continuous. The first row in each panel reports the ex-post alpha (% per month) for various points and percentiles of the performance distribution, ranging from worst fund (min) to best fund (max). The second row reports the associated t-statistic for these alphas. Row 3 contains the ex-post t-statistics of alpha ranked from lowest (min) to highest (max). Row 4 reports the bootstrap p-values of the t-statistics in row 3 based on 1,000 bootstrap resamples. Both actual ex-post and bootstrap t-statistics are based on Newey-West heteroscedasticity and autocorrelation adjusted standard errors. The final row reports the number of monthly observations for the fund. The full sample period is 1975M4:2002M12 (333 observations).

All Investment Objectives : Unconditional Three Factor Model : $(R_i - r_f)_t = \alpha_t + \beta_{1i}(R_m - r_f)_t + \beta_{2i}SML_t + \beta_{3i}HML_t + \varepsilon_{it}$																		
Panel A : Minimum Fund History of 18 observations																		
	min	5.min	min5%	min10%	min40%	max40%	max10%	max5%	max3%	20max	15max	12max	10max	7max	5.max	3.max	2.max	max
Alpha	-3.057	-1.847	-0.498	-0.386	-1.115	-0.016	0.293	0.491	0.616	0.756	1.022	1.085	1.119	1.371	1.447	2.138	2.212	2.234
t-alpha	-1.990	-3.615	-2.194	-1.098	-1.420	-0.106	1.627	0.758	1.212	1.108	1.420	3.083	1.013	2.926	3.365	4.530	2.800	2.225
t-stat	-7.414	-4.724	-2.971	-2.450	-0.790	-0.106	1.390	1.901	2.387	2.671	2.796	2.926	2.991	3.365	3.639	4.023	4.282	4.530
p-tstat	0.002	0.000	0.000	0.000	0.000	1.000	0.182	0.013	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.006	0.020	0.122
No. Obs	333	333	44	333	41	134	152	56	23	70	37	23	63	48	34	164	31	23
Panel B : Minimum Fund History of 36 observations																		
	min	5.min	min5%	min10%	min40%	max40%	max10%	max5%	max3%	20max	15max	12max	10max	7max	5.max	3.max	2.max	max
Alpha	-0.901	-0.666	-0.431	-0.345	-0.108	-0.019	0.225	0.377	0.490	0.530	0.588	0.598	0.616	0.686	0.723	0.756	1.447	2.234
t-alpha	-2.532	-7.414	-1.577	-1.294	-2.450	-0.152	1.498	2.529	2.677	1.698	4.023	2.556	1.212	2.991	1.322	1.901	3.365	2.225
t-stat	-7.414	-4.724	-3.043	-2.508	-0.836	-0.140	1.242	1.719	2.196	2.403	2.544	2.671	2.677	2.796	2.991	3.389	3.710	4.023
p-tstat	0.000	0.000	0.000	0.000	0.000	1.000	0.915	0.395	0.002	0.001	0.003	0.004	0.012	0.049	0.067	0.031	0.039	0.123
No. Obs	333	333	95	186	139	228	116	94	137	185	178	70	38	37	63	212	42	164
Panel C : Minimum Fund History of 60 observations																		
	min	5.min	min5%	min10%	min40%	max40%	max10%	max5%	max3%	20max	15max	12max	10max	7max	5.max	3.max	2.max	max
Alpha	-0.901	-0.663	-0.427	-0.330	-0.108	-0.027	0.181	0.301	0.411	0.439	0.478	0.507	0.530	0.588	0.598	0.686	0.723	0.745
t-alpha	-2.532	-4.445	-1.563	-2.406	-1.591	-0.235	1.757	1.920	3.389	0.806	2.777	2.776	1.698	4.023	2.556	2.991	1.322	2.546
t-stat	-7.414	-4.724	-3.077	-2.537	-0.958	-0.216	1.242	1.670	2.023	2.196	2.403	2.522	2.544	2.671	2.776	2.991	3.389	4.023
p-tstat	0.000	0.000	0.000	0.000	0.000	1.000	0.824	0.508	0.169	0.016	0.004	0.006	0.022	0.066	0.132	0.159	0.059	0.056
No. Obs.	333	333	137	79	197	228	116	279	148	137	185	207	178	70	159	63	212	164

'Mutual Fund Performance: Skill or Luck ?'

Panel D : Minimum Fund History of 120 observations																		
	min	5.min	min5%	min10%	min40%	max40%	max10%	max5%	max3%	20max	15max	12max	10max	7max	5.max	3.max	2.max	max
Alpha	-0.673	-0.650	-0.426	-0.345	-0.118	-0.036	0.144	0.240	0.317	0.284	0.319	0.382	0.395	0.439	0.507	0.588	0.606	0.745
t-alpha	-2.050	-2.970	-2.708	-3.318	-0.440	-0.406	0.839	2.522	1.610	2.673	1.927	2.403	1.859	0.806	2.776	4.023	2.292	2.546
t-stat	-7.414	-4.724	-3.146	-2.608	-1.019	-0.318	1.129	1.610	1.927	1.757	2.023	2.245	2.292	2.529	2.546	2.776	3.389	4.023
p-tstat	0.000	0.000	0.000	0.000	0.000	1.000	0.993	0.173	0.442	0.674	0.238	0.059	0.105	0.060	0.242	0.239	0.026	0.027
No. Obs	333	333	128	140	171	275	170	127	333	218	148	207	227	129	179	159	212	164

Figure 1 : Cross-Sectional Alpha : Unconditional Model

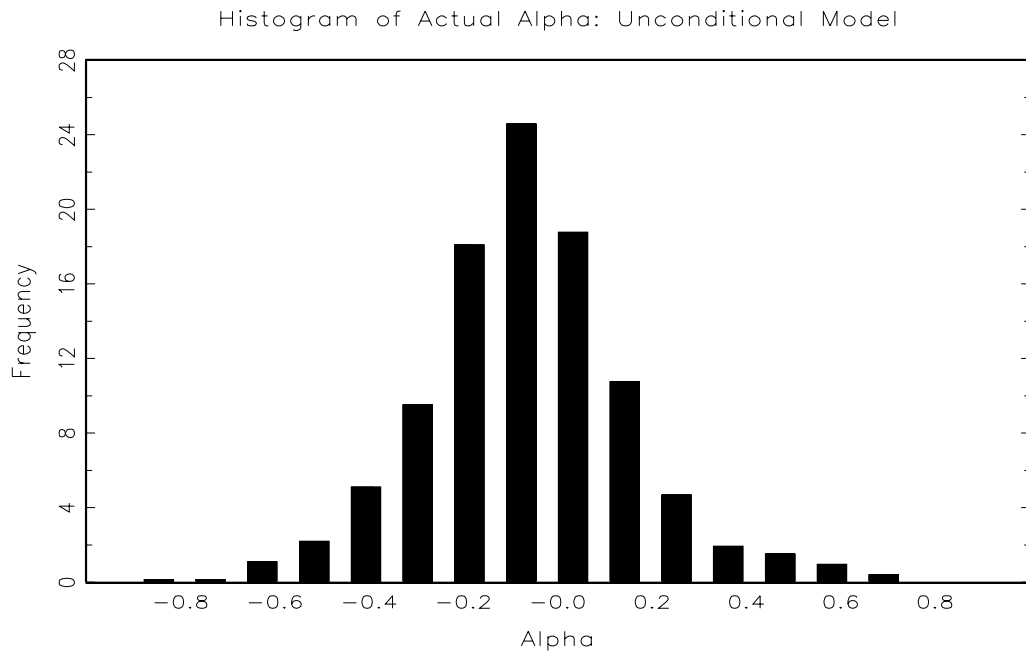


Figure 2 : Cross-Sectional Alpha : Conditional-Beta Model

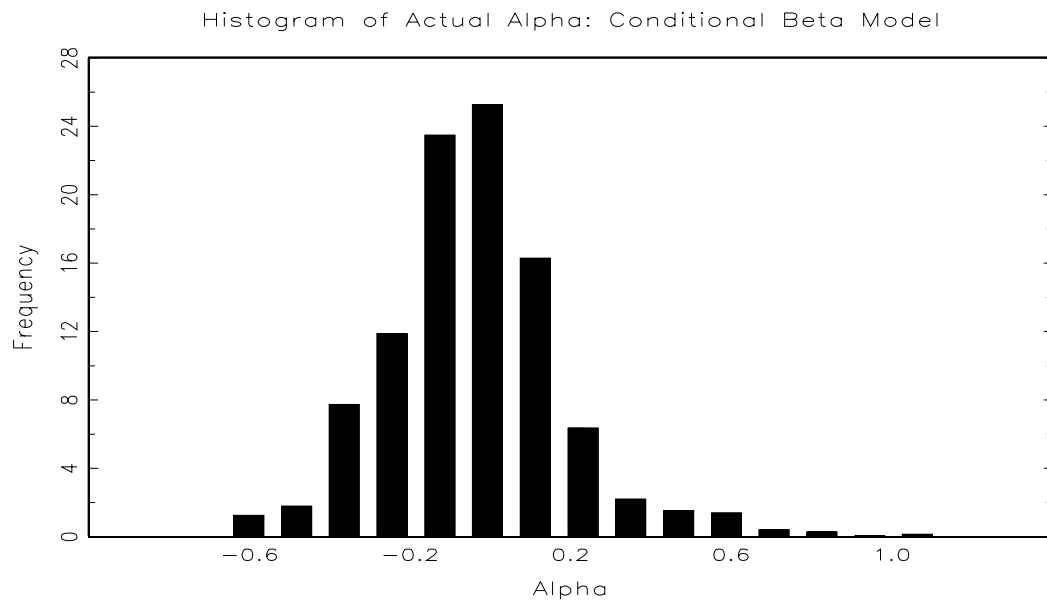
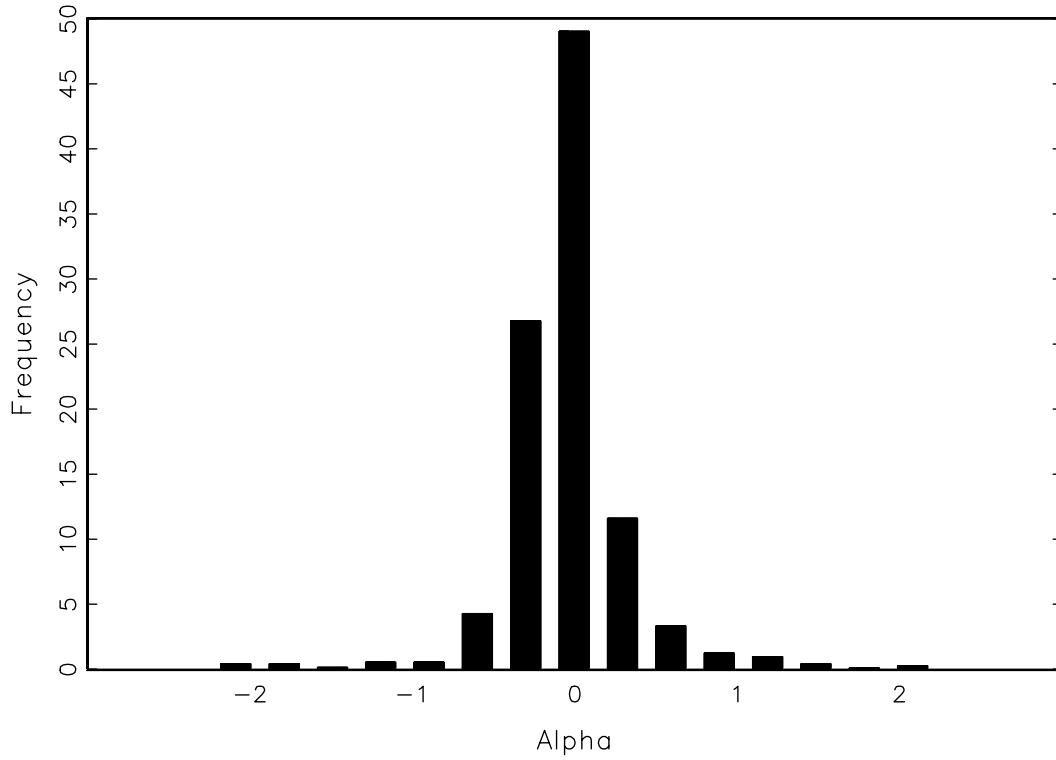


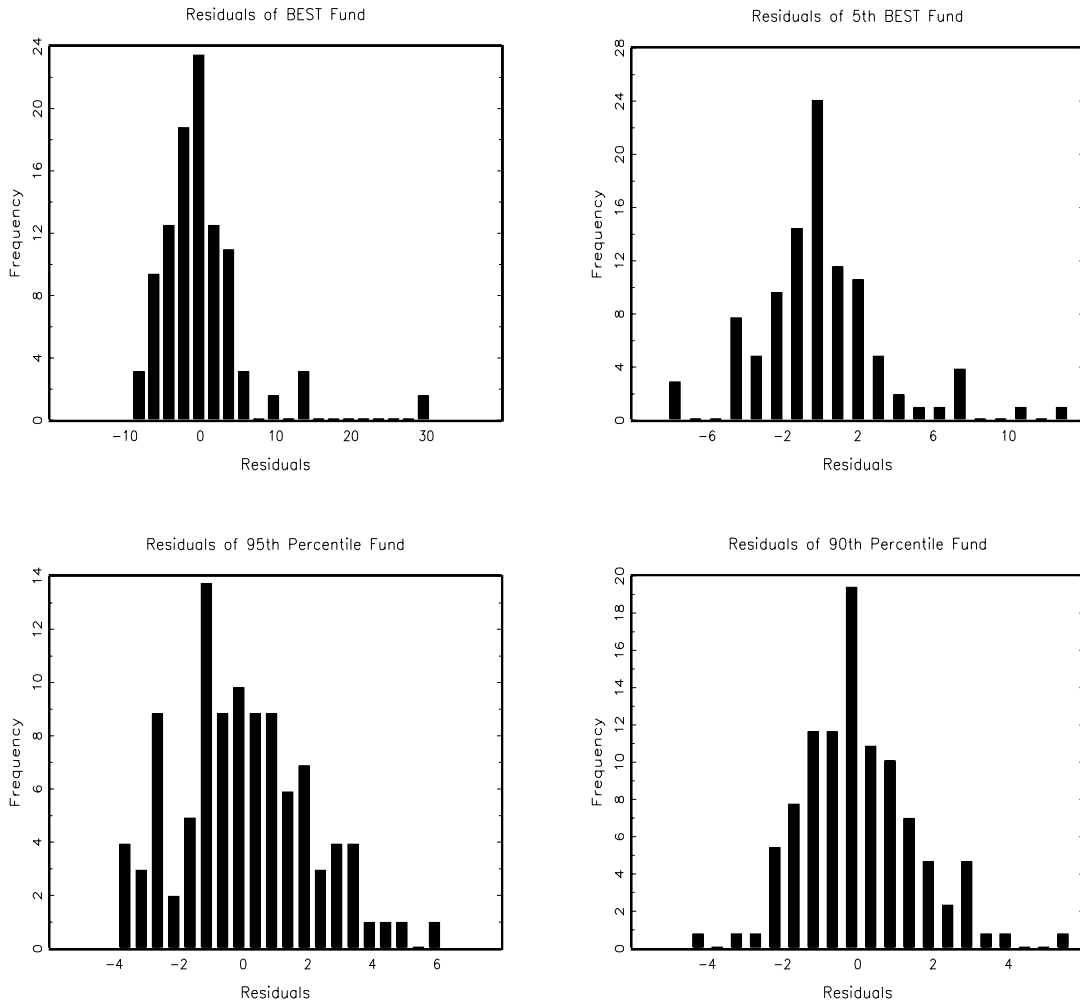
Figure 3 : Cross-Sectional Alpha : Conditional Alpha-Beta Model

Histogram of Actual Alpha: Conditional Alpha and Beta Model



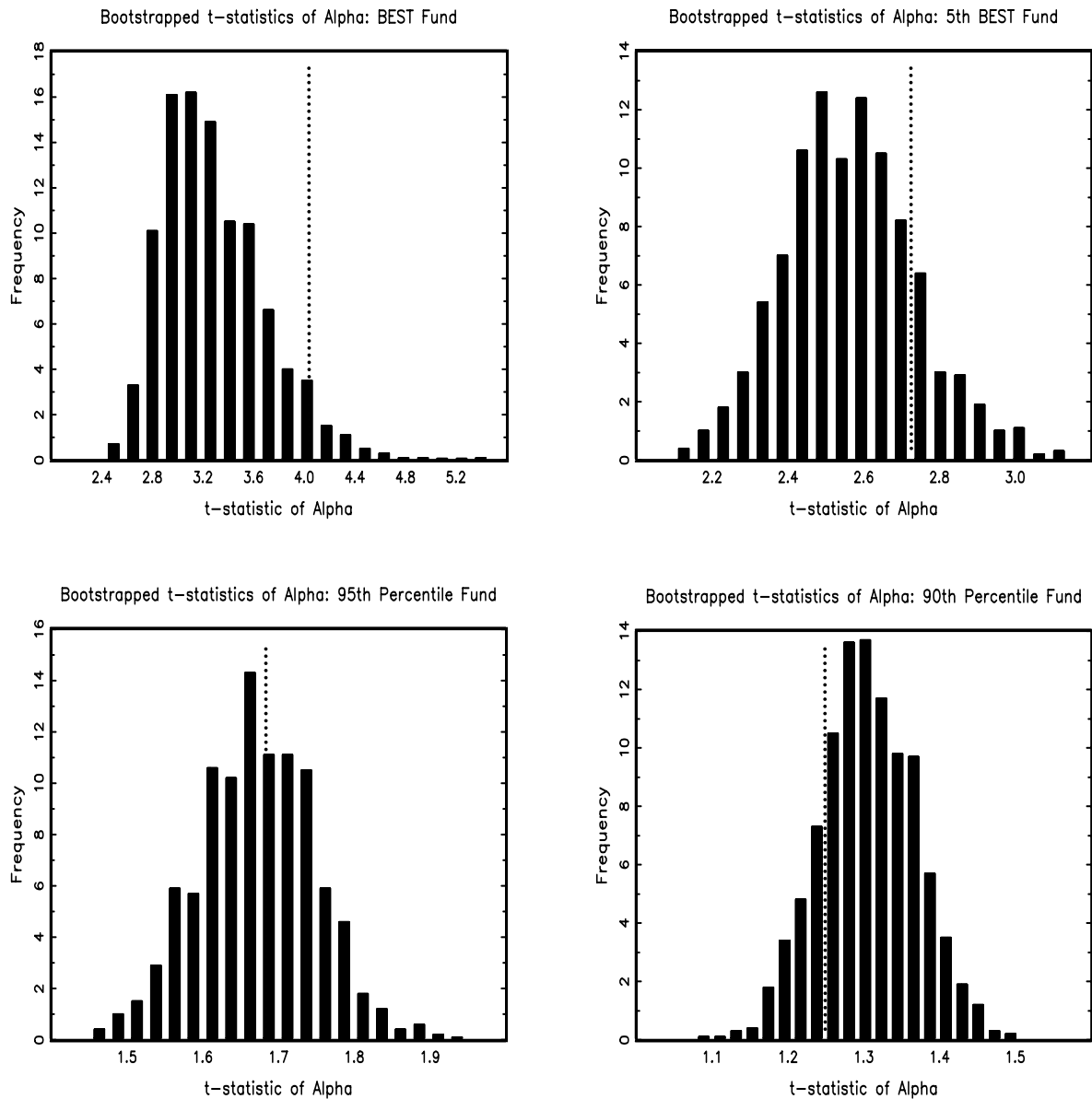
**Figure 4a : Histograms of Residuals
(Upper End of the Distribution)**

Figure 4a shows histograms of the residuals from the estimation of the unconditional FF model, at various points in the upper end of the cross-sectional performance distribution.



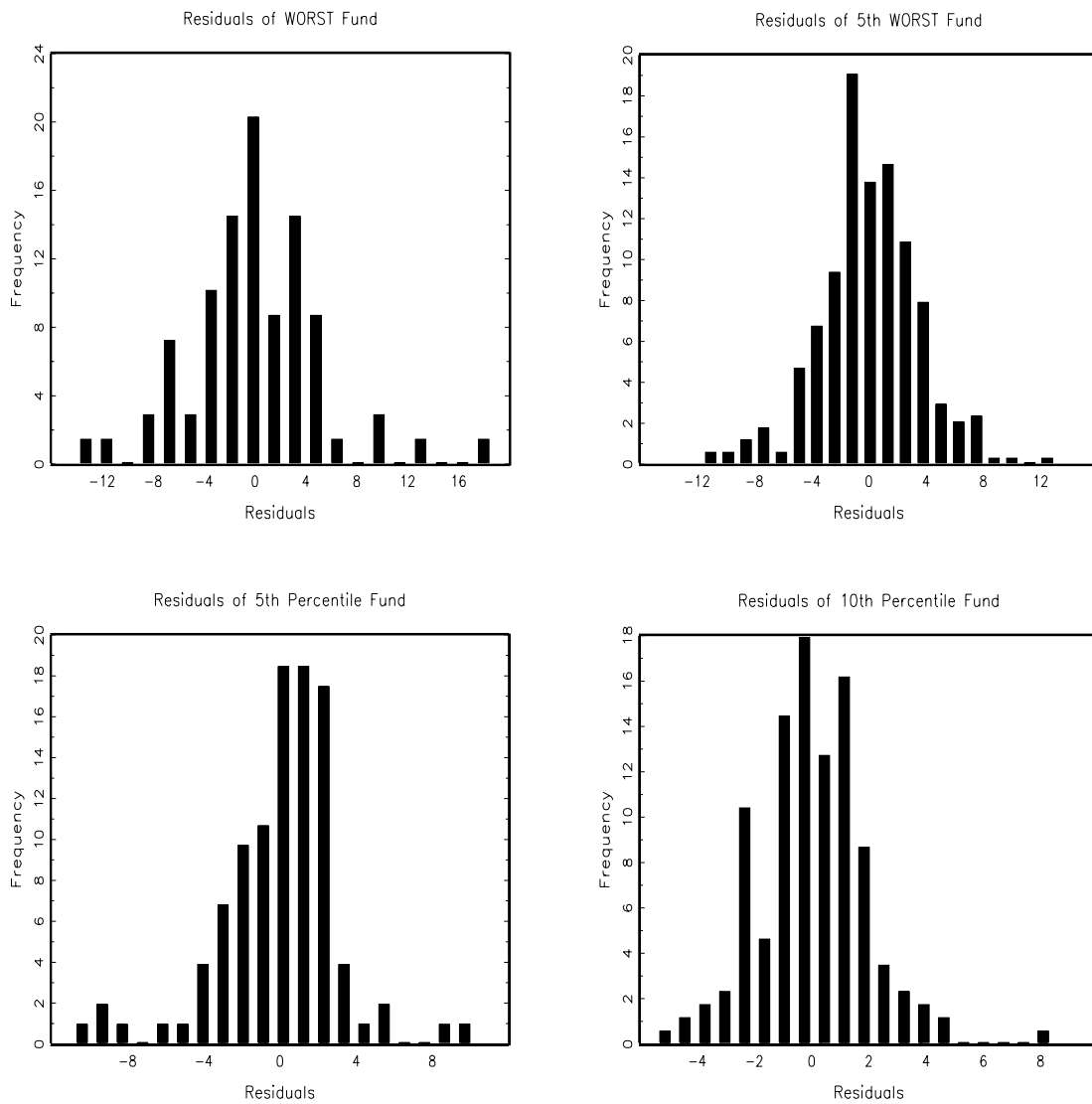
**Figure 4b : Histograms of Bootstrap t-alpha Estimates
(Upper End of the Distribution)**

Figure 4b shows histograms of the bootstrap t-statistics of alpha from the estimation of the unconditional FF model, at various points in the upper end of the performance distribution. The actual t-statistic for alpha is indicated by the vertical dashed line.



**Figure 5a : Histograms of Residuals
(Lower End of the Distribution)**

Figure 5a shows histograms of the residuals from the estimation of the unconditional FF model, at various points in the lower end of the performance distribution.



**Figure 5b. Histograms of Bootstrap t-alpha Estimates
(Lower End of the Distribution)**

Figure 5b shows histograms of the bootstrap t-statistics of alpha from estimation of the unconditional FF models, at various points in the lower end of the performance distribution. The actual t-statistic for alpha is indicated by the vertical dashed line.

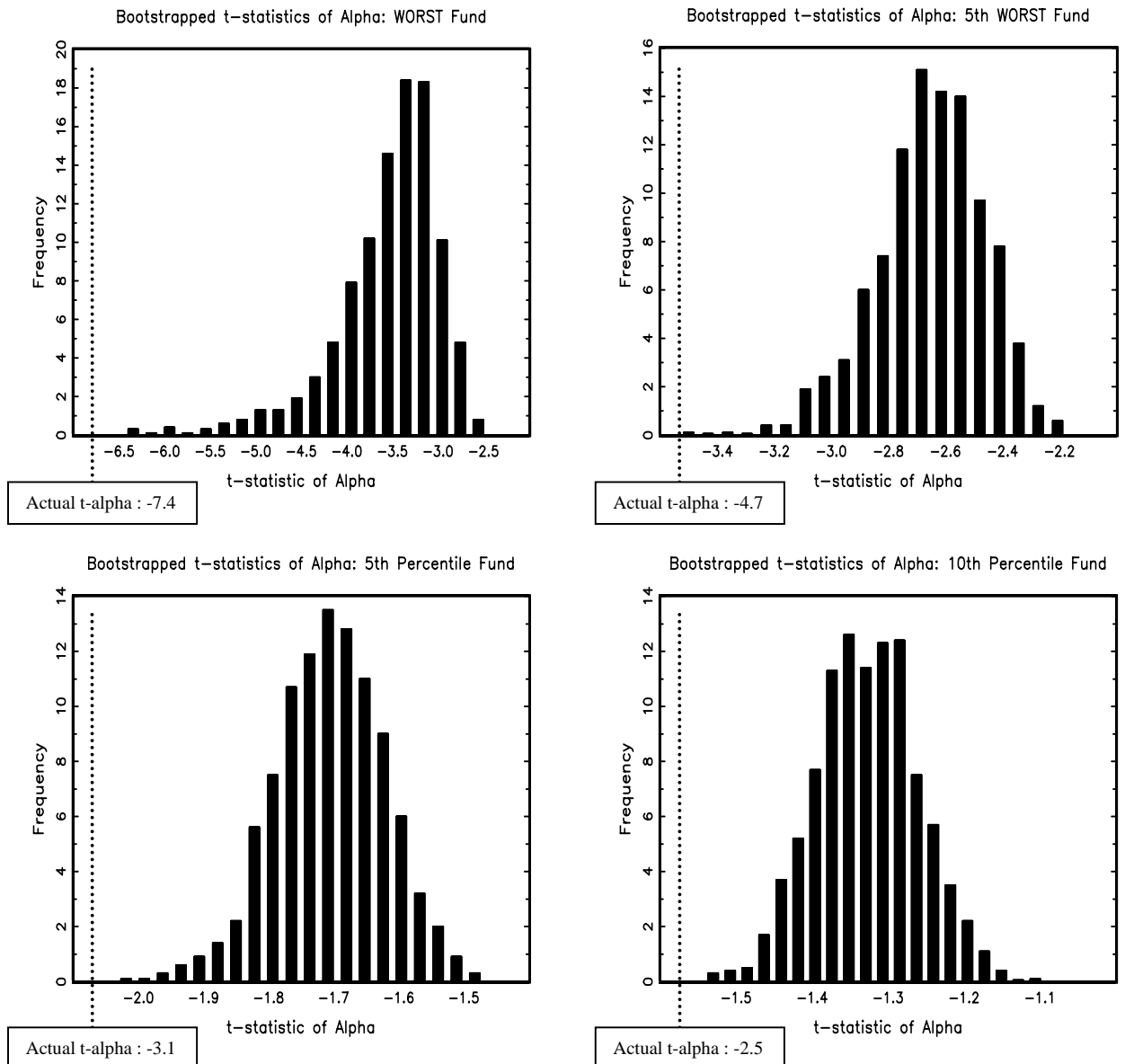


Figure 6. Kernel Density Estimates of the Actual and Bootstrap distribution.

Figure 6 shows the Kernel density for the actual and bootstrap distributions of the t-statistics of alpha from the unconditional FF model over the full sample of mutual funds. Funds are required to have a minimum of 60 observations and t-statistics are Newey-West adjusted. The plots are generated using a Gaussian Kernel.

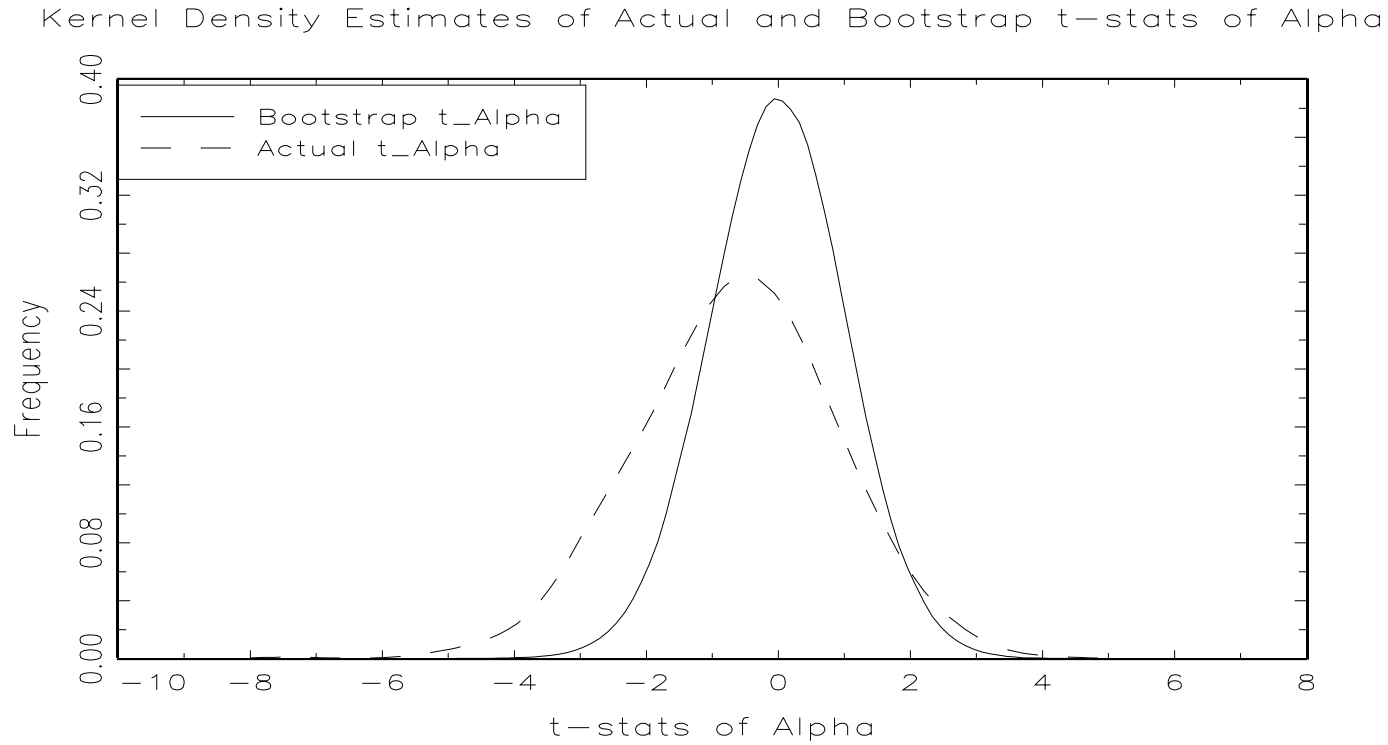


Figure 7. Kernel Density Estimates of the Actual and Bootstrap distribution - by Investment Style

Figure 7 shows the Kernel density for the actual and bootstrap distributions of the t-statistics of alpha using separate bootstraps on the funds of different investment styles. Results relate to the unconditional FF model over the full sample of mutual funds. t-statistics are Newey-West adjusted and funds with a minimum of 60 observations are used. The plots are generated using a Gaussian Kernel.

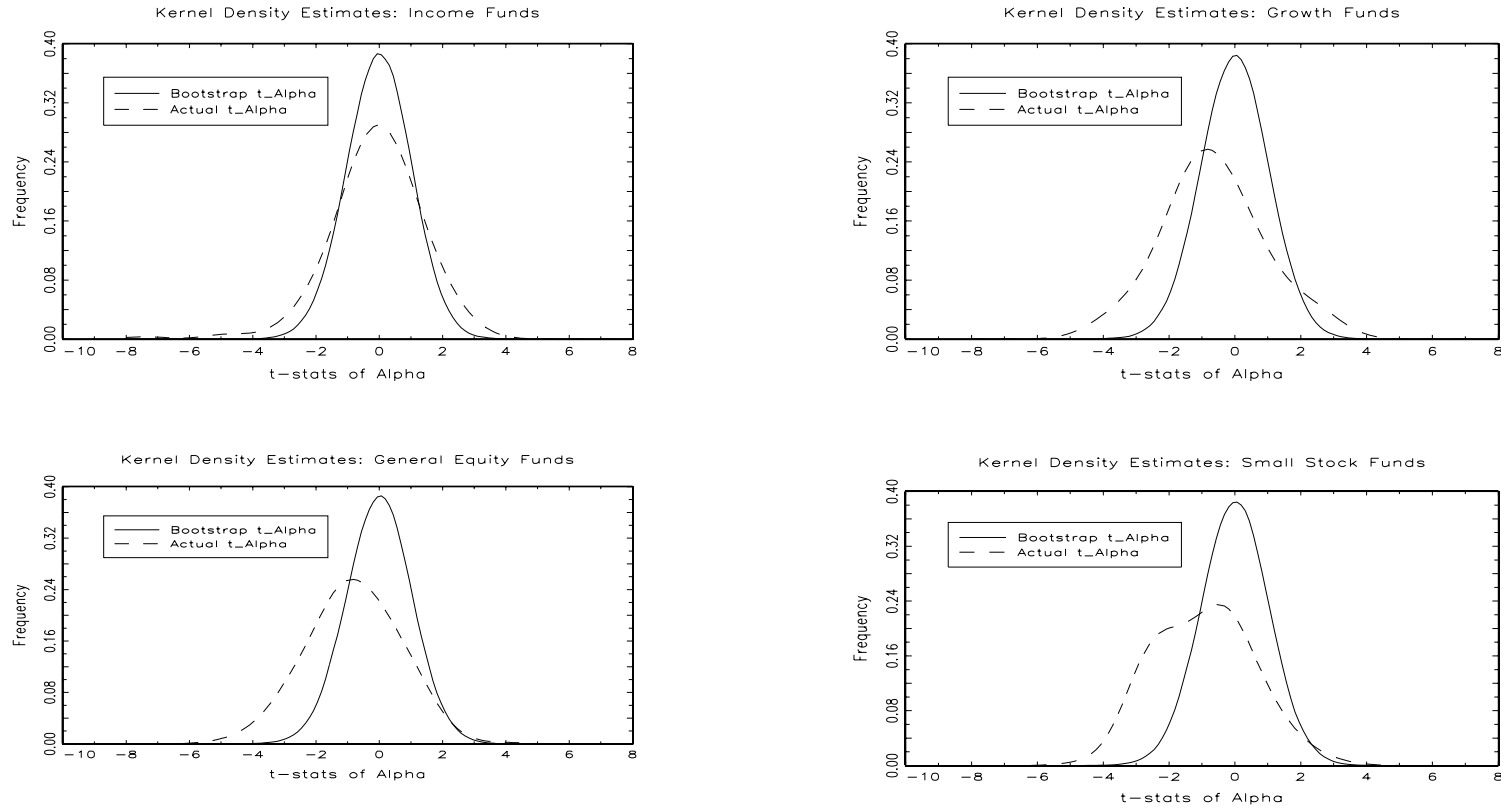


Figure 8. Kernel Density Estimates of the Actual and Bootstrap Distribution - by Location

Figure 8 shows the Kernel density for the actual and bootstrap distributions of the t-statistics of alpha using separate bootstraps on the onshore and offshore funds. Estimates are from the unconditional FF model, t-statistics are Newey-West adjusted and funds with a minimum of 60 observations are used. The plots are generated using a Gaussian Kernel.

