First draft 11/11/04. This version : 4/7/05

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 equity 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 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 *relatively 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 most poor performing funds are merely unlucky. Most of these funds demonstrate 'bad skill'. Recursive estimation and Kalman 'smoothed' coefficients indicate temporal stability in the performance alpha's of winner and loser portfolios.

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. Niall O' Sullivan is grateful for financial assistance from the Arts Faculty Research Fund at University College Cork. We gratefully acknowledge the provision of mutual fund return data by Fenchurch Corporate Services using Standard & Poor's Analytical Software and Data. Main programmes use GAUSS™.
2. We thank, Don Bredin, Stuart Hyde, David Miles, Turalay Kenc, Ales Cerny, Aneel Keswani, Ian Marsh, Gulnar Muradoglu, Andrew Sykes (FSA), Dylan Thomas, Giovanni Urga and participants at Smirfit Business School, Dublin, University of Bordeaux IV, University of Bari, MMF Conference London 2004, International Finance Conference-University College Cork 2004, Global Finance Conference-Trinity College, Dublin 2005.

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, Baks et al 2001, Pastor and Stambaugh 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. 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 2003, 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 2000). Following on from this Wermers (2003) finds that persistent large cash inflows to winner funds are invested with a lag and 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 under-perform (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

December 2002. A data set of over 900 equity 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 return predictability is present (Brennan et al 1997, Campbell et al 2003) or, when there is uncertainty about parameters in the prediction equation for returns (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% of their assets 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 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 1996, 2000, Hon and Tonks 2003), 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, 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 and helps mitigate survival bias problems (Brown, Goetzmann, Ibbotson and Ross 1992). 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 236 'nonsurviving' funds in the analysis.

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, 7 funds exhibit levels of performance which cannot be attributable to 'luck' at 5% significance level, while 12 funds exhibit such performance at 10% significance level. As we move further towards the centre of the performance distribution (e.g. below the 97% percentile) many 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 poorly performing active funds in the universe of UK equity 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 inertia, learning and search costs when trying to evaluate alternatives.

When examining different fund 'styles', we find genuine outperformance among the top equity income funds but there is little evidence of skill for the top performers amongst the 'all company' and small stock funds. For 'all companies' and small stock funds the extreme left tail of the performance distribution indicates 'bad skill' rather than bad luck – but for income funds the converse applies – the poor performance of income funds is due to bad luck rather than 'bad skill'. 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. In the left tails of these distributions, we find that extreme poor performers (negative alphas), whether they are onshore or offshore, demonstrate 'bad skill' rather than bad luck.

Broadly speaking, 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 survivorship 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'* 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 935 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, they can only be traded between the investor and the trust manager and 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 UK

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

In our database of 935 funds, we remove '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. Furthermore, 93 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 842 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 income (162 funds), 'all companies' (i.e. formerly general equity and equity growth, 553 funds) and smaller companies (127 funds). The data set includes both surviving funds (699) and nonsurviving funds (236). 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). In addition, funds are also categorized by the location of operation. Onshore funds (731) are managed in the UK while offshore funds (204) 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 and net of management fees. Because our focus is on the performance of the fund's managers rather than on returns to investors/customers, our returns data is calculated bid-price to bid-price (with income reinvested).

The market factor used is the FT All Share Index of total returns (i.e. including reinvested 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 Companies (HGSC) Index and the returns on the FT 100 index[1]. 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 index[2]. The factor

[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

mimicking portfolio's momentum behavior, MOM, has been constructed using the constituents of the London Share Price Data Base, (total return) index[3].

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

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

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

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

where SMB_t, HML_t and MOM_t are factor mimicking portfolios for size, book-to-market value and 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

---

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 six months are calculated. The MOM variable is constructed by taking the difference between these two variables. The universe of stocks is the London Share Price Data Base.

the CAPM. In addition, Carhart(1997) finds that momentum is statistically significant in explaining (decile) returns on US mutual funds – although the latter variable is less prevalent in studies on UK data (e.g. Blake and Timmermann 1998, Quigley and Sinquefield 2000, Tonks 2004).

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 over-pricing (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) $$r_{i,t+1} = \alpha_i + b_{0i}(r_{b,t+1}) + B_i^{'}(z_t * r_{b,t+1}) + \varepsilon_{i,t+1}$$

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: α_i = 0.

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

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? (see Admati *et al* 1986). In the model of Treynor and Mazuy (1966) a successful market timer adjusts the market factor loading $\beta_{it} = \theta_i + \gamma_{im}[r_{m,t}]$ so that (1) may now be written:

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

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

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

where $\left[r_{m,t}\right]^+ = \max\{0, r_{m,t}\}$ 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

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 Sinquefield 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 existing 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. H_0: α_i = 0 for i = 1, 2, …, n) each fund's 'true' alpha is normally distributed and each fund has a different but *known* standard deviation 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* α_i = 0 (here i = 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 α_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', with α_i = 0 across all funds, we have $(\alpha_1^{(1)}, \alpha_2^{(1)}, \ldots \ \alpha_n^{(1)})$ 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 $\alpha_{\max}^{(1)}$ 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, …, B trials) we can obtain the complete distribution of $\alpha_{\max}$ under the null of no outperformance, which we denote $f(\alpha_{\max})$.

Note that the distribution $f(\alpha_{\max})$ 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 $\hat{\alpha}_{\max}$ against the 'luck distribution' for the best fund. If $\hat{\alpha}_{\max}$ exceeds the 5% right tail cut off point in $f(\alpha_{\max})$, we can reject the null hypothesis that the performance of the best fund is attributable to luck.

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.

The above demonstrates the main features and intuition behind our analysis of fund performance - which is based on the theory of order statistics. 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 under the null $f(\alpha_{\max})$, encapsulates all of the different individual fund's 'luck distributions' (and in a multivariate context this cannot be derived analytically from the theory of order statistics).

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 empirical '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 fund[4]. 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 (see Politis and Romano 1994). Consider an estimated model of equilibrium returns of the form:

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

for i = {1, 2, …, n) funds, where T_i = number of observations on fund-i, $r_{i,t}$ = $\left(R_{i,t} - R_{f,t}\right)$, 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 vectors $\left\{\hat{\beta}_i, e_{i,t}\right\}$. Next, for each fund-*i*, draw a random sample (with replacement) of length T_i from the 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 α_i = 0):

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

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, …, 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

[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.

bootstrap simulations. The bootstrap estimates of $\tilde{\alpha}_i$ may be gathered in a matrix of dimension (n x B) as follows.

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

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 α_i = 0. This is the 'luck distribution' for the extreme top performer, $f(\tilde{\alpha}_{\max})$. 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 $f(\tilde{\alpha}_i)$, 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 *ex-post* $\hat{\alpha}_i$ 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 from $f(\tilde{\alpha}_{\max})$ then we reject the null that its performance is due to luck (at 95% confidence). We infer that the fund 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.

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' $\sigma^2_{\varepsilon_i}$ 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 $T_{i,\min}$ = 36 months to minimize survivorship bias.

4. Model Selection

In this section, the equilibrium models described in section 2 are examined. 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 $T_{i,\min}$ = 36, to minimize survivorship bias. For each model, cross-sectional (across funds) average statistics are calculated. A single 'best model' is chosen from each of the 3 model classes; (i) unconditional, (ii) conditional-beta and (iii) conditional alpha-beta, and these results are reported in table 1. (In all, we examined over 50 models within the three classes of model and these results are available on request). In none of our models are the Treynor-Mazuy (1966) and Merton-Henriksson (1981) market timing variables significant. The key model selection metrics are the Schwartz Information Criterion (SIC) and the statistical significance of the individual parameters.

[Table 1 here]

In the best three models, the cross-sectional average alpha takes on a small and statistically insignificant negative value. 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 table 1, where we find that the alpha of an *equally weighted* portfolio of all funds have statistically significant negative alphas (for all three models).

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 0.26% p.m. (3.1% p.a.), for the unconditional and conditional-beta models and somewhat larger at 0.75%p.m. for the conditional alpha-beta model. 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 holding 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 three classes of model, whereas the HML factor beta is often not statistically significant, even at a 10% significance level (as discussed further at the end of the next section). We find that the momentum factor (MOM) is generally not statistically significant at the individual UK fund level (e.g. Blake and Timmermann 1998, Tonks 2004), in contrast to US studies (Carhart 1997). For the conditional-beta model (2nd column, table 1) only the dividend yield variable produces near statistically significant results. In the conditional alpha-beta model we find that none of the conditional alphas has a t-statistic greater than 1.1 but some of the conditional betas are bordering on statistical significance and our best model is shown in column 3.

The above results suggest that the unconditional Fama-French 3 factor model explains UK equity mutual fund returns data reasonably well. There is little additional explanatory power from the conditional and market timing variables (not reported). The latter 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 table 1), the adjusted R^2 across all three models is around 0.8, while around 64% of funds have non-normal errors (Bera-Jarque statistic), and around 40% of funds have serial correlation (which is of order one – LM statistic). The Schwartz Information Criterion (SIC) is lowest for the unconditional model.

The Fama-French 3 factor model was 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 models 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 36 observations for a fund to be included in the analysis. This leaves a sample of 675 funds, of which 189 are non-survivor funds (i.e. have ceased to exist at some point before the end of the sample period), while 486 are survivor funds.

In Table 1, we reported that around 64% 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. Figure 4 shows the distribution of the residuals for selected funds in the upper (i.e. 'best' and 90th percentile fund) and lower tail (i.e. 'worst' and 10th percentile fund) of the (ex-post) performance distribution. Residuals from funds in the extreme tails (e.g. 'best' and 'worst' funds) tend to exhibit higher variance and a greater degree of non-normality than residuals from funds closer to the centre of the performance distribution (e.g. 90th and 10th percentiles).

For funds in the upper tail, it is this high variance coupled with large positive residuals, that causes these 'best funds' to populate the very top end of the bootstrap distributions. In turn this generates wide dispersion and non-normality in the performance of the very 'top performing' funds. This is evident in figure 5a which shows the bootstrap histograms of $t_{\tilde{\alpha}_i}$ at selected points of the performance distribution. Figure 5b presents 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 each of the tails do not, and it is the latter in which investors are particularly interested.

[Figure 4 here]
[Figures 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 each fund's actual (ex-post) 't-alpha', ranked from lowest to highest (left to right) and the second row shows its associated value of 'alpha'. Row 3 ("p-tstat") reports the bootstrap p-values of the ranked t-statistics in row 1, based on the 'luck distribution' for $t_{\tilde{\alpha}_i}$ under the null of no outperformance.

[Table 2 here]

For example, using the *unconditional* model the 'max' fund (Table 2, Panel A) has an actual ex-post $t_{\hat{\alpha}}$ = 3.389 and achieved an abnormal performance of $\hat{\alpha}_{\max}$ = 0.412% p.m. However, the bootstrap p-value of t-alpha for the 'max' fund is 0.437 (row 4). The latter indicates that from among the 1,000 bootstrap simulations across all funds, under the null hypothesis of

zero abnormal performance, 43.7% of the bootstrap t-statistics for the highest ranked fund were greater than $t_{\hat{\alpha}}$ = 3.389. This can be seen in the histogram in top left of figure 4b, where the vertical line shows the actual $t_{\hat{\alpha}}$ = 3.389, relative to the 'luck distribution'. Thus using a 5% upper tail cut off point, we cannot reject the hypothesis that the best fund's actual $t_{\hat{\alpha}}$ = 3.389 may be explained by luck alone. Thus while the conventional $t_{\hat{\alpha}}$ = 3.389 of the best fund indicates genuine skill, the non-parametric bootstrap indicates 'good luck'. This apparent contradiction is due to the highly non-normal distribution of idiosyncratic risk across our top performing funds in the right tail of the performance distribution. It demonstrates that standard test statistics may give misleading inferences when we look at funds in the extreme tails – as can be seen for example, for funds up to '7 max' in table 2, panel A.

Our complete set of bootstrap results show that of the top 20 ranked funds, 12 achieve genuine outperformance at a 10% significance level while 7 funds outperform at a 5% significance level. However, as one moves into the centre of the performance distribution (i.e. at or below the top 3% of funds) there is no evidence of stock picking ability – the bootstrap indicates that any positive $t_{\hat{\alpha}}$'s are due to luck rather than skill (see table 2, panel A and figure 4b).

In the left tail of the distribution, (i.e. the left side of Panel A, table 2), the lowest ranked fund has $t_{\hat{\alpha}}$ = –5.358 with a bootstrap p-value of 0.009. Hence for the ex-post 'worst fund' 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 $t_{\hat{\alpha}}$ = –5.358, which is to the left of the 'luck distribution'. It is clear from the left hand side of 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 10 funds to achieve $\hat{\alpha}$ ≥ 0.5% p.m. (6% p.a.) based on random chance alone, whereas 19 funds exhibit this level of performance (or higher). However, $\hat{\alpha}$ ≥ 0.1% p.m. (1.2% p.a.) is expected to be achieved by 173 funds solely based on chance, while in fact only 142 funds are observed to have reached this level of performance. Of course, this interpretation is consistent with the discussion of p-values

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 points 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 tail of the distribution of actual $t_{\hat{\alpha}}$'s 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 'bad skill'. In contrast, the extreme right tail of the distribution of $t_{\hat{\alpha}}$ for the 'real data' lies outside the 'luck distribution'. This means there are some, but not many, genuine 'outperformers'.

Panels B and C of Table 2 reports findings from the conditional-beta and conditional alpha-beta FF models. Inferences from the bootstrap (rows 't-alpha', 'p-tstat'), for the left tail of the performance distribution are very similar to those for the unconditional FF model in Panel A - 'bad luck' is again not a defense for bad performance.

.

The results for the right tail of the distribution using the two conditional FF models (Panels B and C, Table 2) are broadly consistent with those for the unconditional model (Panel A). Genuine stock picking ability is found for around 7% of funds using the condition-beta model and for about 4% of funds using the conditional alpha-beta model, but it is luck rather than skill which accounts for the positive performance of many funds further towards the centre of the performance distribution.

In other results (available on request) we find that the removal of the HML_t variable produces virtually no changes from those reported in table 2, while addition of the momentum variable produces slightly more winners (around 5%) than the unconditional 3 factor model (of table 1, Panel A). These models also support the view that many poorly performing funds have bad skill rather than bad luck.

Our results are qualitatively consistent with Kosowski et al (2004) who find strong evidence of stock picking ability among top performing 5-10% of US funds (depending on the model chosen) and genuine poor performance for the funds in most of the left tail of the performance distribution.

Above we applied the bootstrap across all funds using each of our 3 'best models'. However, recall from Table 1 that the set of conditioning information variables were shown to be only weakly statistically determined (on average across funds) 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 3 factor model of panel A and our variants (described below) use this 'baseline model'.

Performance and Investment Styles

Having found some 'good skill' and lots of 'bad skill' when analyzing all UK mutual funds together, the question now arises whether these 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 675 funds have a minimum of 36 monthly observations of which 143 (21%) are income funds, 423 (63%) are 'all companies' funds and 109 (16%) are small stock funds. To further address the 'style question' we apply the bootstrap procedure separately for each style, since the distribution of idiosyncratic risk may differ across different styles.

[Table 3 here]

Table 3, Panels A, B and C re-estimate the performance statistics of table 2, for the three investment styles. Looking at the left side of all three Panels in Table 3 '(t-alpha', 'p-tstat') it is clear that genuine 'bad skill' in the left tail is common across 'all companies' and small stock funds, whereas poorly performing income funds experience bad luck rather than 'bad skill'.

Looking at the right side of all three panels of Table 3, it is mainly high ranking equity income funds (Panel A) which achieve positive levels of performance, which cannot be accounted for by luck. In particular, we find that most equity income funds ranked from the 3rd highest to the 'max 10%' generally beat the bootstrap estimate of luck (at a 5% significance level), while the performance of the two highest ranked income funds could have been achieved by luck alone. In contrast to the above, for UK 'All Companies' and small stocks (Table 3, Panels B and C, 't-alphat', 'p-tstat'), there are hardly any funds which have genuine stock picking skills in the right tail of the performance distribution. Note that amongst these top performing funds, standard t-

tests would often give different inferences to the bootstrap (e.g. see the 'max', '2 max' and '3 max' funds for equity income and 'UK all companies').

The above results are consistent with those in table 2, where of the 7 funds with genuine skill (at a 5% significance level), 6 can be identified as income funds and one as a small company fund, whereas at a 10% significance level we have 12 skilled funds of which 6 are income funds, 5 are 'all companies' and 1 is a small company fund. Hence, in table 2 income funds are proportionately more representative of skill, than the other fund styles.

Our findings for the UK of 'skill' mainly among some top performing UK *income* funds are in contrast to results in Kosowski et al (2004) for US mutual funds, who find that it is the top performing *growth* funds that have genuine skill. (But note that Kosowski et al 2004 do not have a 'small companies' style classification).

[Figure 7 here]

Figure 7 shows Kernel density estimates for the three investment styles. For equity income 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 the other two style categories, actual ex-post performance does not exceed random sampling variation[5]. We also see that the left tails of the actual ex-post t-statistics for all companies and small companies, lie largely to the left of the 'luck distributions', indicating that poor performance is unlikely to be due to bad luck.

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.

[5] 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.

[Table 4 here]

[Figure 8 here]

The bootstrap results in table 4 ('t-alpha', 'p-tstat') are clear cut. Out of 553 onshore funds almost all of the top 20 possess genuine skill (panel A), whereas any positive abnormal performance by the 122 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. These results are also consistent with those of table 2, where all of 12 'skilled' funds (FF unconditional model, 10% significance level) can be identified as onshore funds.

Performance and Survival

Some nonsurvivor funds are represented among funds whose performance is superior to chance. For example, among the top 20 ranked funds (using the unconditional FF model, panel A, table 2), 7 funds beat luck at 5% significance level, 2 of which are nonsurvivors while 12 funds beat luck at a 10% significance level, 3 of which are nonsurvivors.

A possible explanation for the positive performance of nonsurviving funds is that some of these 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 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 their sample period. A large number of such 'mergers' may be due to good rather than bad performance - this is broadly consistent with behavior in a competitive funds market in the theoretical model of Berk and Green (2004).

For the unconditional FF model (panel A, table 2) we find many poorly performing funds with 'bad skill'. 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.

Extensions of the Bootstrap

The 'baseline' bootstrap procedure described in section 3 can be modified to incorporate further characteristics of the data, for example, serial correlation in residuals or, independent residual and factor resampling or, allowing for contemporaneous correlation 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 our inferences reported above, regarding skill versus luck in performance, are very similar.

FUND OF FUNDS

Using the bootstrap (on t-alpha) and the complete sample period, we have identified a few funds that exhibit out-performance that is not due solely to luck, and many funds whose poor performance is due to 'bad skill' rather than bad luck. Following on from this, a natural question is whether a *portfolio* of the 'best' or 'worse' performing funds have constant parameters and in particular, constant *alphas.* Here, we are not looking at portfolio performance from an *ex-ante* viewpoint (as in studies of performance persistence) but examining whether the long-run risk adjusted performance of a chosen portfolio of funds has constant parameters, as we move through time. Such evidence would complement our bootstrap results which use 't-alpha' (from the whole sample) as the performance criterion. Investors want to be reasonably sure that any 'alpha-performance' is not sample specific.

To investigate this issue we apply recursive OLS (with GMM correction for standard errors) as well as the Kalman filter to the portfolio parameters - using the unconditional Fama-French 3 factor model. Recursive estimation allows the parameters to slowly change over time as new data becomes available. The Kalman filter random coefficients model has the parameters $\beta_{k,t}$ on the market return, the SMB and HML factors and the portfolio 'alpha' α_t follow:

(9) $$\beta_{k,t} = \beta_{k,t-1} + v_{k,t} \qquad k = 1,2,3.$$

(10) $$\alpha_t = \alpha_{t-1} + v_t$$

This specification allows the α_t , $\beta_{k,t}$ to change considerably from period-to-period depending on the variances of the error terms σ_ν , which we assume are contemporaneously uncorrelated. (Recursive OLS is a special case of the Kalman filter where $\sigma_\nu = 0$ - see Hamilton 1994).

For an equally weighted portfolio of the best 12 funds (as identified in the bootstrap), the recursive OLS estimates (with GMM corrected standard errors) are shown in figure 9. The recursive market beta coefficient is remarkably constant at around 0.9, as is the factor loading on SMB which is constant at around 0.25, while the factor loading on HML is not statistically different from zero for much of the time period. The 12 best funds from the bootstrap also appear to give a genuine constant out-performance as the recursive estimate of alpha is around 0.58 (6.96% p.a.) over the whole recursive period to end December 2002.

[Figure 9 and 10 here]

At the bottom end of the performance distribution (figure 10) the key feature of the equally weighted portfolio of the worst 50 funds is the near constancy of the negative alpha of around -0.35 (-4.2% p.a.). Like the 'best' funds, the factor loadings on the market beta and SMB are reasonably constant, while that on HML is again statistically insignificant. This general qualitative pattern is repeated when we include all of the worst funds up to the 40th percentile in our equally weighted portfolio – all of these funds have 'bad skill' as indicated by our bootstrap procedure.

[Table 5 here]

The Kalman filter estimates of the one-step ahead ($\alpha_{t|t-1}$) and smoothed alphas($\alpha_{T|t}$) for the 'best' and 'worse' portfolios are similar to those for recursive OLS discussed above so in table 5 we present the final state vectors (at time T), their p-values and the standard error (σ_ν) of the time varying parameters (see Hamilton 1994). A portfolio of the best 12 funds as indicated by the bootstrap analysis (which *all* have actual t-alphas significant at a 10% critical value) has a final state vector α_T = 0.446 (p-value = 0.0083). The relatively low standard error σ_ν = 0.0279 of the time varying alpha confirms that our ranking on bootstrap t-alphas does provide a portfolio of 'top' funds that exhibits constant outperformance over the whole period of around 5.3% p.a. - before taking account of transactions costs (but after payment of management fees). The standard error

of the time varying market beta is virtually zero indicating this factor loading is constant. As with the recursive OLS results, the HML factor is not statistically different from zero.

For a portfolio of the worst 50 funds, alpha varies little (σ_ν = 0.0003) and the final state vector is α_T = -0.337 (p-value < 0.0001) – similar to the recursive OLS results. Therefore a constant poor performance of around -4.0% p.a. can be attributed to many of the worst performing funds, even when we allow for the possibility of time varying parameters.

6. Conclusion

Using a comprehensive data set for the UK, April 1975 – December 2002, with over 900 equity mutual funds, we use a bootstrap methodology to distinguish between 'skill' and 'luck'. Depending on the particular model chosen, we find genuine stock picking ability for somewhere between 5 and 10 percent of top performing UK equity mutual funds (i.e. performance which is not solely due to good luck). This is consistent with recent US empirical evidence (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.

Controlling for different investment objectives among funds, it is found that some of the top ranked equity income funds show genuine stock picking skills whereas such ability is generally not found among small stock funds and 'all company' funds. We also find that positive performance amongst onshore funds is due to genuine skill, whereas for offshore funds, positive performance is attributable to luck.

Note that the above results do not necessarily imply that the mutual fund industry is inefficient, since in a competitive market we only expect a few funds to earn positive risk adjusted returns over long horizons. This is because funds with genuine skill and high past returns have large inflows (as observed in many empirical studies) and with increasing marginal costs to active management, this leads to zero long run average profits for most funds (Berk and Green 2004).

At the negative end of the performance scale, our *whole analysis* strongly rejects the hypothesis that most poor performing funds are merely unlucky. Most of these funds demonstrate 'bad skill', which is broadly consistent with results for US funds (Kosowski et al 2004). This result is not consistent with the competitive model of Berk and Green (2004), since 'bad skill' should lead to large outflows of funds and the return on funds who survive should, in equilibrium, equal that on a passive (index) fund. 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), indicates that many investors either cannot correctly evaluate fund performance or find it 'costly' to switch between funds or suffer from a disposition effect.

Our bootstrap ranks funds on the basis of t-alpha using the whole sample of data. However, when we apply recursive OLS or the Kalman filter-random coefficients model to a *portfolio* of funds based on our bootstrap rankings, we find considerable stability in the estimated portfolio alphas (as well as the market return and SMB factor loadings). This suggests that there is genuine constant outperformance amongst a few top funds and genuine underperformance amongst many poorly performing funds.

For the active fund management industry as a whole, our findings are something of a curate's egg. For the majority of funds with positive abnormal performance, we find this can be attributed to 'good luck'. We also show that '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 *individual* active funds which demonstrate genuine skill, based on their track records.

The above results suggest that at the present time many UK equity investors would be better-off holding index/tracker funds, with their lower transactions costs. On the policy side the UK government wishes to encourage long term saving via mutual (and pension) funds (Turner 2004). So perhaps it is time the Financial Services Authority changed its 'health warning' on investing in equity funds from, 'The value of your investments can go down as well as up' to, 'Active fund management may damage your financial health'. Of course, the latter is 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

Admati, A.R., Bhattacharya, S., Ross, S.A. and Pfleiderer, P. (1986) 'On Timing and Selectivity' *Journal of Finance,* 41, 715-730.

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, Elsivier 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-institue.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.

Bollen, N.P.B. and Busse, J.A. (2005) 'Short-Term Persistence in Mutual Fund 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.

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

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.

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. (1992) 'The Persistence of Mutual Fund Performance' Journal of Finance, XLVII(5), 1977-1984.

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

Hamilton, J.D. (1994) *Time Series Analysis*, Princeton University Press, Princeton, N.J.

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.

Hon, M. and Tonks, I. (2003) 'Momentum in the UK Stock Market', *Journal of Multinational Financial Management,* 13(1), 43-70.

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.

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.

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

Leger, L. (1997) 'UK Investment Trusts : Performance, Timing and Selectivity', *Applied Economics Letters*, 4, 207-210.

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 Stambaugh, 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 Sinquefield, 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, forthcoming *Journal of Business 2005.*

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 2 using all mutual funds. Only the best model from each of the 3 classes of model (1) unconditional model (2) conditional-beta and (3) conditional alpha-beta are reported. 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 (Bera-Jarque 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 and 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.

Model	**1**	**2**	**3**
	unconditional	**conditional beta**	**conditional alpha and beta**
Regression Coefficients			
Average alpha (percent p. m.)	-0.057	-0.032	-0.109
Standard Deviation of Alpha	0.261	0.261	0.754
Unconditional Betas (t-stats in parentheses)			
r_{mt}	0.912	0.863	0.849
	(25.196)	(21.193)	(21.068)
SMB_t	0.288	0.285	0.257
	(4.959)	(4.905)	(4.043)
HML_t	-0.025	-0.023	0.016
	(1.451)	(1.451)	(1.207)
Conditional Variables, Z_{t-1} (Dividend Yield)			
$Z_{t-1} * r_{mt}$	-	-0.048	-0.055
		(1.408)	(1.496)
$Z_{t-1} * SMB_t$	-	-	-0.002
			(1.513)
$Z_{t-1} * HML_t$	-	-	0.033
			(1.044)
Z_{t-1}	-	-	-0.073
			(1.037)
Model Selection Criteria			
Adjusted R-squared	0.807	0.811	0.818
SIC	1.352	1.365	1.432
Residuals not normally distributed (% of funds)	64	64	63
LM(1) statistics (% of funds)	40	40	44
LM(6) statistics (% of funds)	34	35	39
Equally weighted Portfolio			
Alpha	-0.139	-0.112	-0.107
t-statistics	-3.051	-2.464	-2.517

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 t-statistics of 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 alpha for these t-statistics. Row 3 reports the bootstrap p-values of the t-statistics 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 36 observations.

Panel A : Unconditional Model

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

	Min	5 min	min5%	min10%	min40%	max30%	max10%	max5%	max3%	20max	15max	12max	10max	7 max	5 max	3 max	2 max	max
t-alpha	-5.358	-4.278	-3.045	-2.509	-0.873	0.212	1.177	1.698	1.955	2.023	2.282	2.501	2.545	2.678	2.777	2.991	3.365	3.389
Alpha	-0.264	-0.400	-0.165	-0.435	-0.107	0.024	0.216	0.530	0.186	0.447	0.386	0.302	0.431	0.491	0.478	0.686	1.447	0.412
p-tstat	0.009	<0.001	<0.001	<0.001	<0.001	1	0.978	0.491	0.447	0.284	0.070	0.020	0.038	0.094	0.157	0.232	0.128	0.437

Panel B : Conditional-Beta Model

$$(R_i - r_f)_t = \alpha_i + \beta_{1i}(R_m - r_f)_t + \beta_{2i}SMB_t + \beta_{3i}\,HML_t + \beta_{4i}[Z3_{t-1}\,(R_m - r_f)_t] + \varepsilon_{it}$$

	min	5 min	min5%	min10%	min40%	max30%	max10%	max5%	max3%	20max	15max	12max	10max	7 max	5 max	3 max	2 max	max
t-alpha	-5.334	-4.161	-2.929	-2.376	-0.723	0.354	1.394	1.856	2.090	2.189	2.305	2.456	2.622	2.809	3.036	3.161	3.353	4.036
Alpha	-0.262	-0.229	-0.596	-0.222	-0.148	0.064	0.189	0.166	0.801	0.568	2.309	0.661	0.429	0.515	1.431	0.580	0.751	0.449
p-tstat	0.012	<0.001	<0.001	<0.001	<0.001	1	0.147	0.034	0.093	0.022	0.048	0.045	0.017	0.017	0.022	0.093	0.153	0.100

Panel C : Conditional Alpha-Beta Model

$$(R_i - r_f)_t = \alpha_{0i} + \alpha_{1i}Z3_{t-1} + \beta_{1i}(R_m - r_f)_t + \beta_{2i}SMB_t + \beta_{3i}\,HML_t + \beta_{4i}[Z3_{t-1}\,(R_m - r_f)_t] + \beta_{5i}[Z3_{t-1}\,SMB_t] + \beta_{6i}[Z3_{t-1}\,HML_t] + \varepsilon_{it}$$

	min	5 min	min5%	min10%	min40%	max30%	max10%	max5%	max3%	20max	15max	12max	10max	7 max	5 max	3 max	2 max	Max
t-alpha	-9.943	-4.106	-2.870	-2.251	-0.785	0.307	1.313	1.894	2.460	2.499	2.725	2.868	3.064	3.20	3.538	3.732	3.785	5.081
Alpha	-3.545	-0.484	-0.189	-0.738	-0.123	0.237	0.815	0.486	1.063	1.261	1.191	1.125	0.456	0.511	1.480	2.529	1.893	3.802
p-tstat	<0.001	0.001	<0.001	<0.001	<0.001	1	0.881	0.211	0.001	0.002	0.002	0.002	0.002	0.008	0.006	0.055	0.169	0.069

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 (Equity Income funds), Panel B (UK All Companies funds) and Panel C (Smaller Companies funds). All results use the unconditional Fama and French three-factor model. The first row in each panel reports the the ex-post t-statistics of alpha, ranked from lowest (min) to highest (max). The second row reports the associated alpha (% per month) for these t-statistics. Row 3 reports the bootstrap p-values of the t-statistics 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 36 observations.

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

Panel A : Equity Income

	min	5 min	min5%	min10%	min20%	min40%	max30%	max20%	max10%	10max	7 max	5 max	3 max	2 max	Max
t-alpha	-2.954	-2.166	-1.863	-1.488	-0.912	-0.197	0.685	0.977	1.667	1.740	2.245	2.501	2.545	2.634	2.777
Alpha	-0.330	-0.204	-0.314	-0.164	-0.122	-0.028	0.089	0.470	0.184	0.275	0.229	0.302	0.431	0.279	0.478
p-tstat	0.337	0.165	0.179	0.141	0.398	0.721	0.111	0.171	0.007	0.107	0.008	0.005	0.067	0.181	0.415

Panel B : UK All Companies

	min	5 min	min5%	min10%	min20%	min40%	max30%	max20%	max10%	10max	7 max	5 max	3 max	2 max	Max
t-alpha	-5.359	-4.190	-3.118	-2.575	-2.009	-1.084	0.011	0.509	1.045	2.023	2.282	2.672	2.776	2.965	3.389
Alpha	-0.264	-0.355	-0.226	-0.491	-0.164	-0.099	0.009	0.063	0.414	0.447	0.386	0.543	0.507	0.479	0.412
p-tstat	0.014	<0.001	<0.001	<0.001	<0.001	<0.001	1	1	1	0.643	0.422	0.083	0.285	0.296	0.301

Panel C : Smaller Companies

	min	5 min	min5%	min10%	min20%	min40%	max30%	max20%	max10%	10max	7 max	5 max	3 max	2 max	Max
t-alpha	-4.953	-3.117	-3.095	2.772	-2.464	-1.128	0.008	0.397	1.286	1.356	1.610	1.742	2.226	2.991	3.365
Alpha	-0.350	-0.360	-0.479	-0.405	-0.522	-0.278	0.016	0.092	0.253	0.472	0.318	0.716	2.235	0.686	1.447
p-tstat	0.001	<0.001	<0.001	<0.001	<0.001	<0.001	1	1	0.647	0.624	0.490	0.579	0.256	0.040	0.096

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 t-statistics of alpha for various points and percentiles of the performance distribution, ranging from worst fund (min) to best fund (max). The second row reports the associated alpha values (% per month) for these t-statistics. Row 3 reports the bootstrap p-values of the t-statistics 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 36 observations.

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

Panel A : Onshore UK funds

	min	5 min	min5%	min10%	min20%	min40%	max20%	max10%	20max	15max	10max	5 max	3 max	2 max	max
t-alpha	-5.359	-4.119	-2.915	-2.464	-1.681	-0.739	0.755	1.287	2.023	2.282	2.545	2.777	2.991	3.365	3.389
Alpha	-0.264	-0.321	-0.510	-0.522	-0.362	-0.0630	0.158	0.130	0.447	0.386	0.431	0.478	0.686	1.447	0.412
p-tstat	0.011	<0.001	<0.001	<0.001	<0.001	<0.001	0.970	0.615	0.057	0.010	0.006	0.073	0.107	0.063	0.306

Panel B : Offshore

	min	5 min	min5%	min10%	min20%	min40%	max20%	max10%	20max	15max	10max	5 max	3 max	2 max	max
t-alpha	-4.278	-3.751	-3.521	-2.953	-2.330	-1.454	0.008	0.489	0.067	0.429	0.881	1.101	1.390	1.450	1.948
Alpha	-0.400	-0.414	-0.310	-0.440	-0.310	-0.226	0.004	0.075	0.006	0.104	0.355	0.588	0.284	0.593	0.487
p-tstat	0.012	<0.001	<0.001	<0.001	<0.001	<0.001	1	1	1	1	1	1	0.998	1	0.985

Table 5. Kalman Filter : Fund of Funds (July 1985 – December 2002)

The final state vector is the value of $\phi = (\alpha, \beta)$ at the end of the sample period with its associated p-value given in the next column. The standard error of the time varying parameter is the estimate σ_v, where $\phi_t = \phi_{t-1} + v_t$ is the 'state' or 'transition' equation.

		Portfolio of Best 12 Funds			**Portfolio of Worst 50 Funds**		
		Final State Vector	**p-value**	**Standard Error Time Varying Parameter**	**Final State Vector**	**p-value**	**Standard Error Time Varying Parameter**
Constant	α_T	0.4463	0.0083	0.0279	-0.3379	< 0.0001	0.0003
R_m	β_{1T}	0.9077	< 0.0001	< 0.0001	0.9465	< 0.0001	0.0007
SMB	β_{2T}	0.3168	0.0001	0.0189	0.2335	< 0.0001	0.0004
HML	β_{3T}	-0.1241	0.0573	0.0157	-0.0339	0.2192	0.0004
Log Likelihood			-348.33			-249.35	

Figure 1 : Cross-Sectional Alpha : Unconditional Model

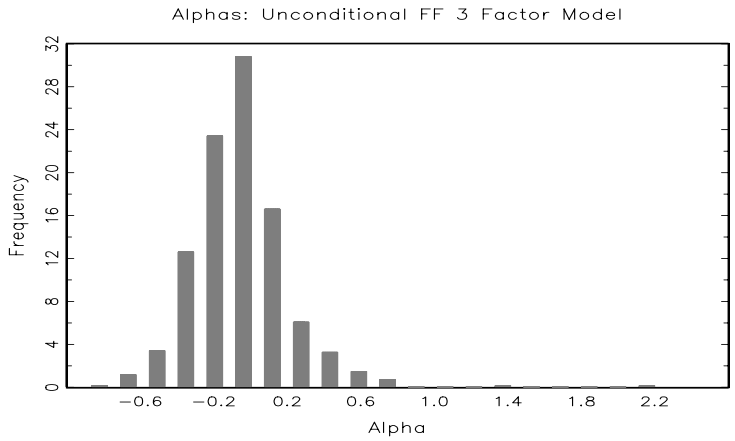
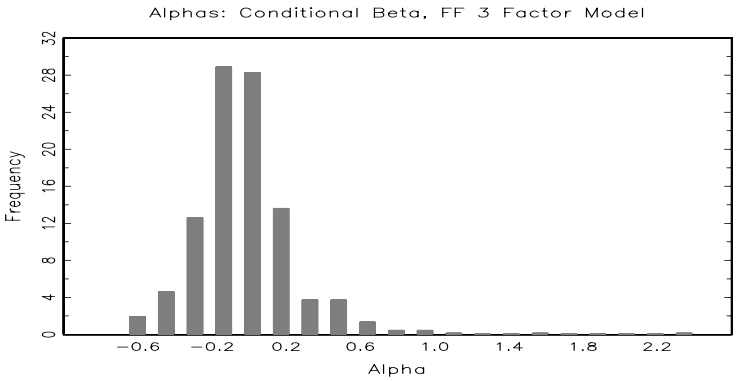


Figure 2 : Cross-Sectional Alpha : Conditional-Beta Model



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

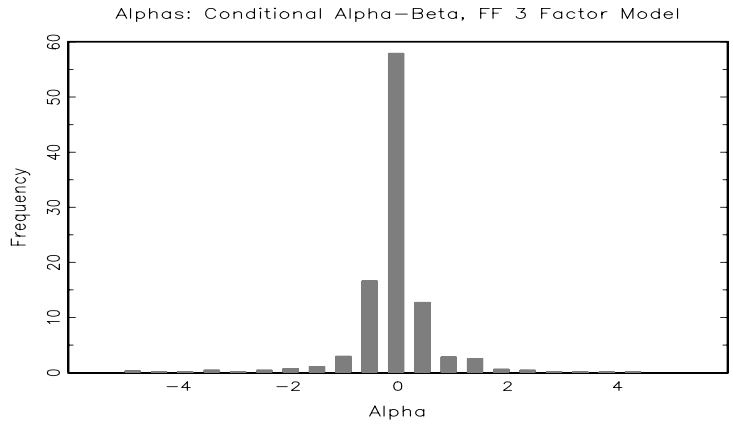


Figure 4 : Histograms of Residuals

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

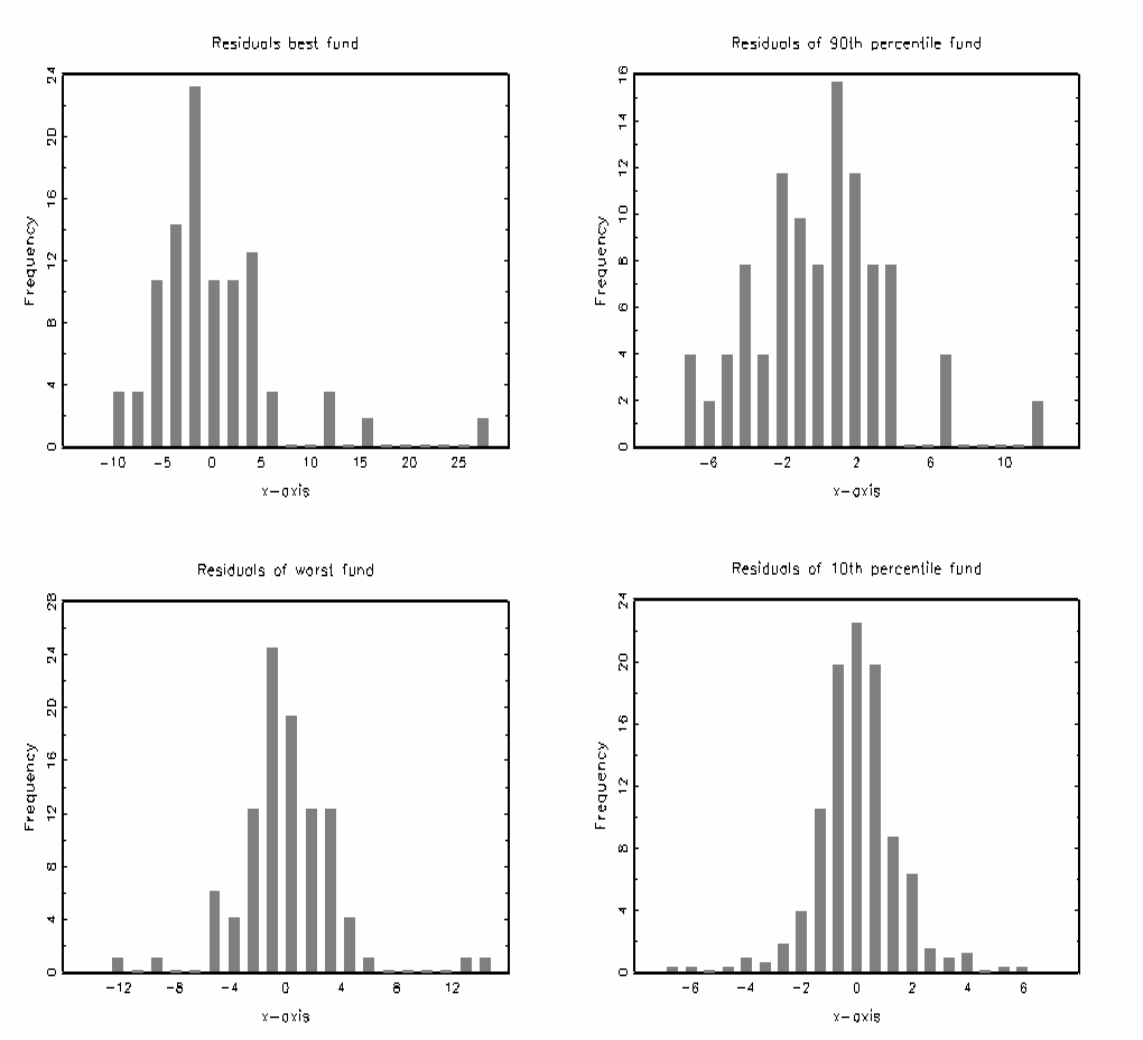


Figure 5a : Histograms of Bootstrap t-alpha Estimates (Upper End of the Distribution)

Figure 5a shows histograms of the bootstrap t-statistics of alpha (under $H_0 : \alpha_i = 0$) at various points in the upper end of the performance distribution (using the 3 factor FF model). The actual (ex-post) t-statistics $t_{\hat{\alpha}}$ are indicated by the vertical dashed line.

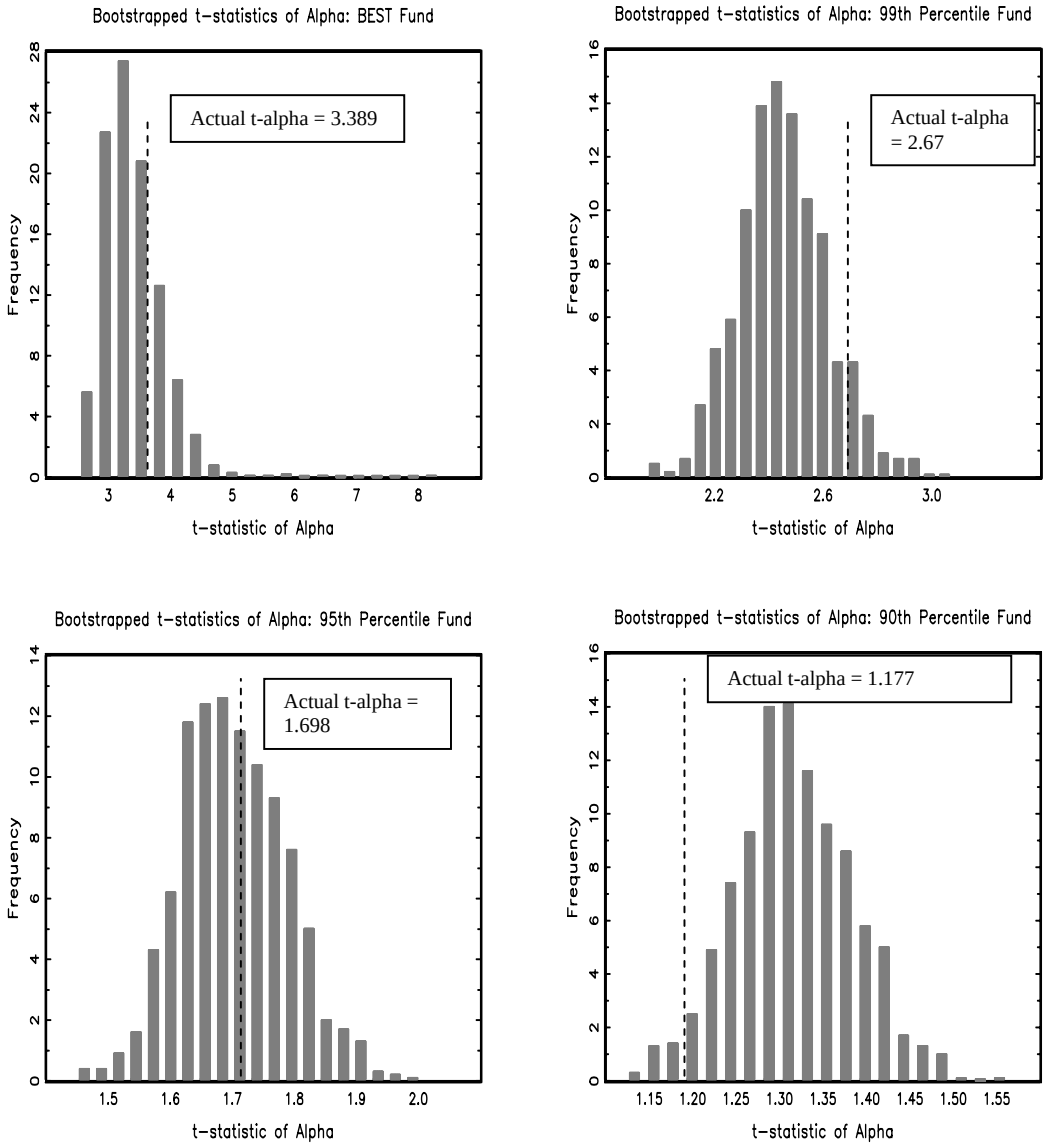


Figure 5b. Histograms of Bootstrap t-alpha Estimates (Lower End of the Distribution)

Figure 5b shows histograms of the bootstrap t-statistics of alpha (under $H_0 : \alpha_i = 0$) at various points in the lower end of the performance distribution (using the 3 factor FF model). The actual (ex-post) t-statistics $t_{\hat{\alpha}}$ are indicated by the vertical dashed line.

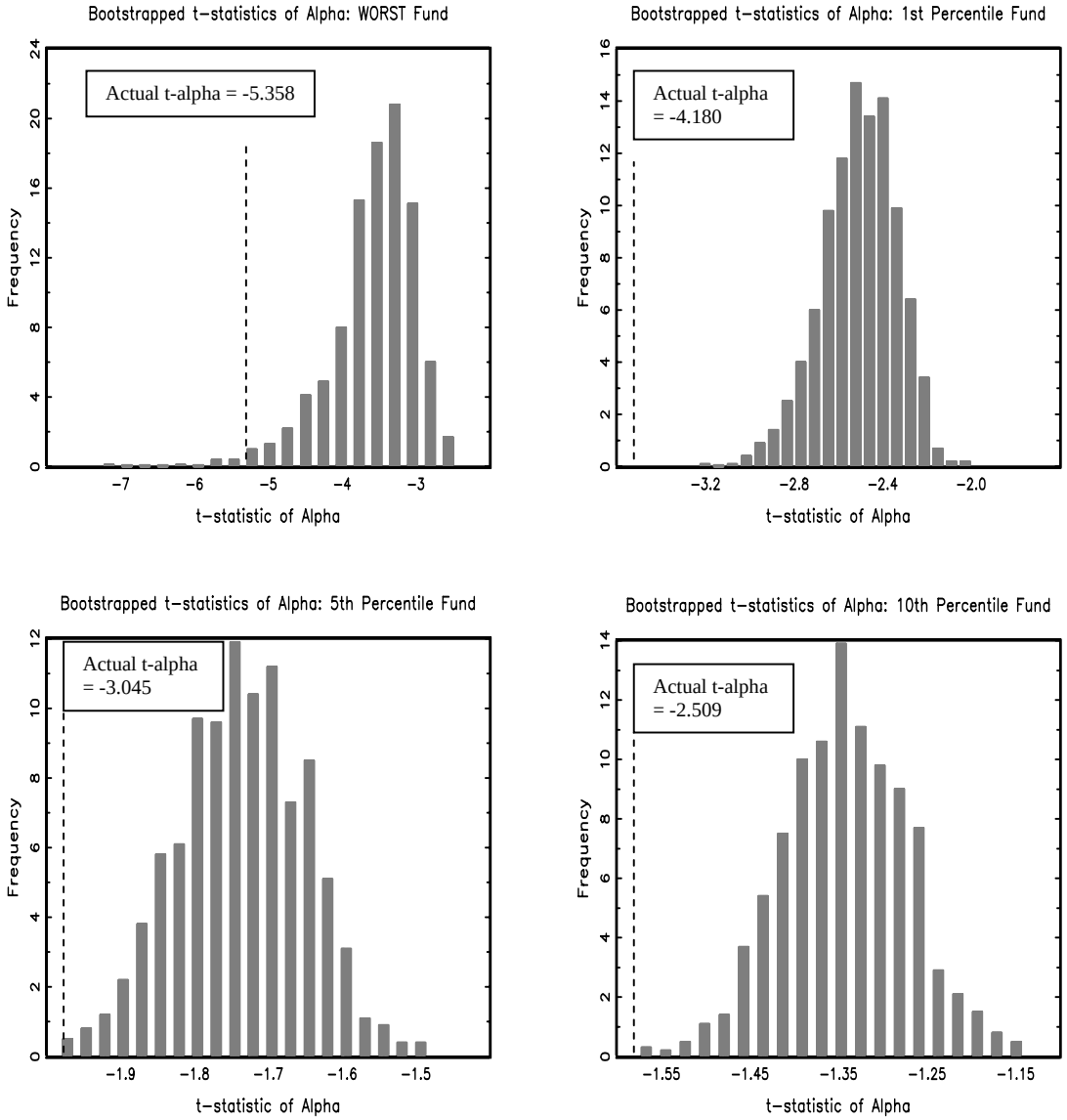


Figure 6. Kernel Density Estimates of the Actual and Bootstrap distribution of t-alphas – All Funds.

Figure 6 shows Kernel densities of 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 36 observations and t-statistics are Newey-West adjusted. The plots are generated using a Gaussian Kernel.

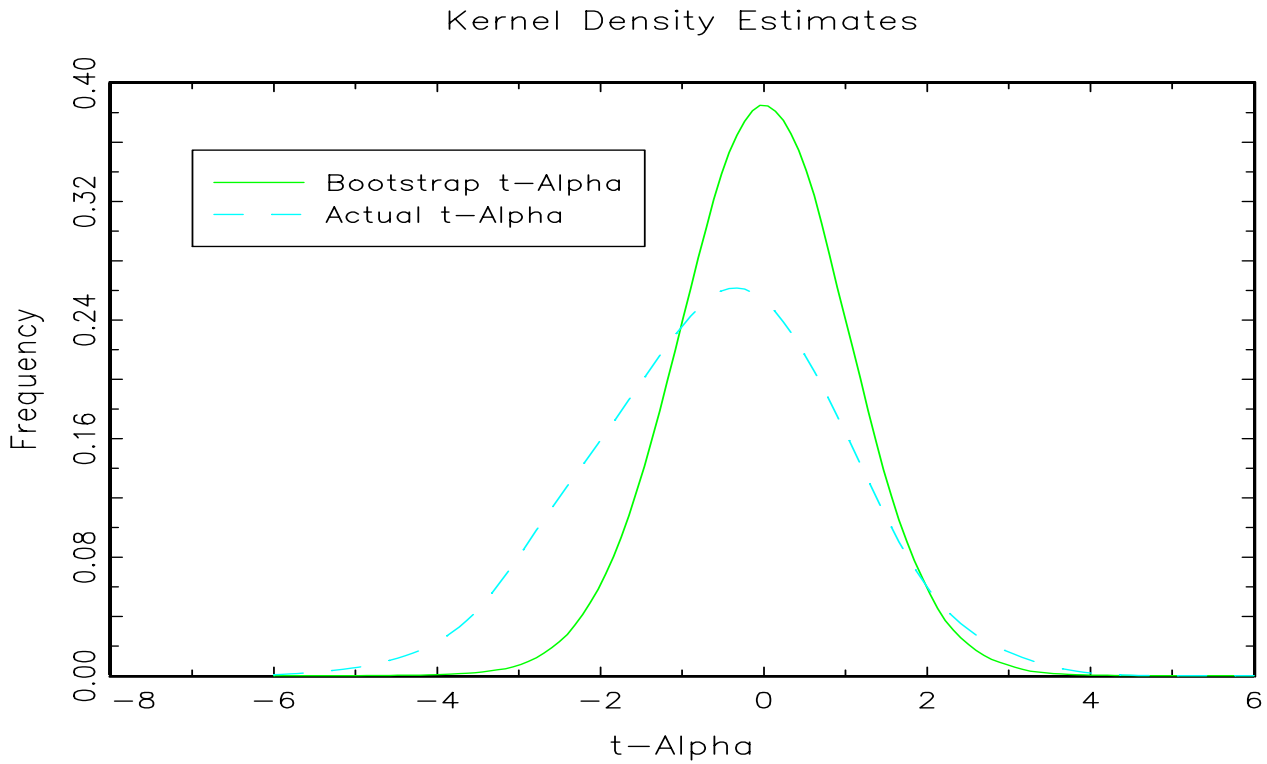
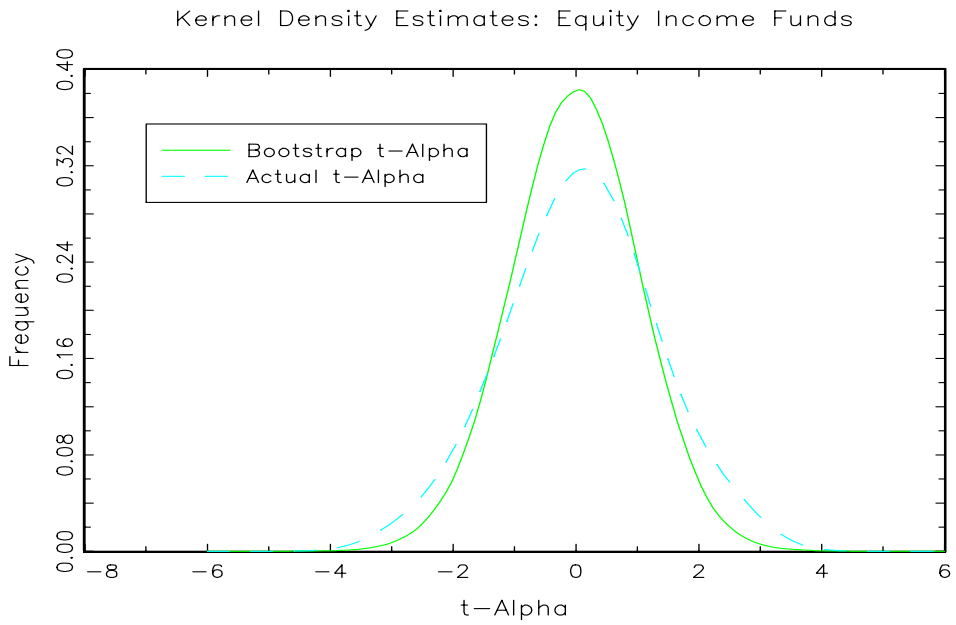


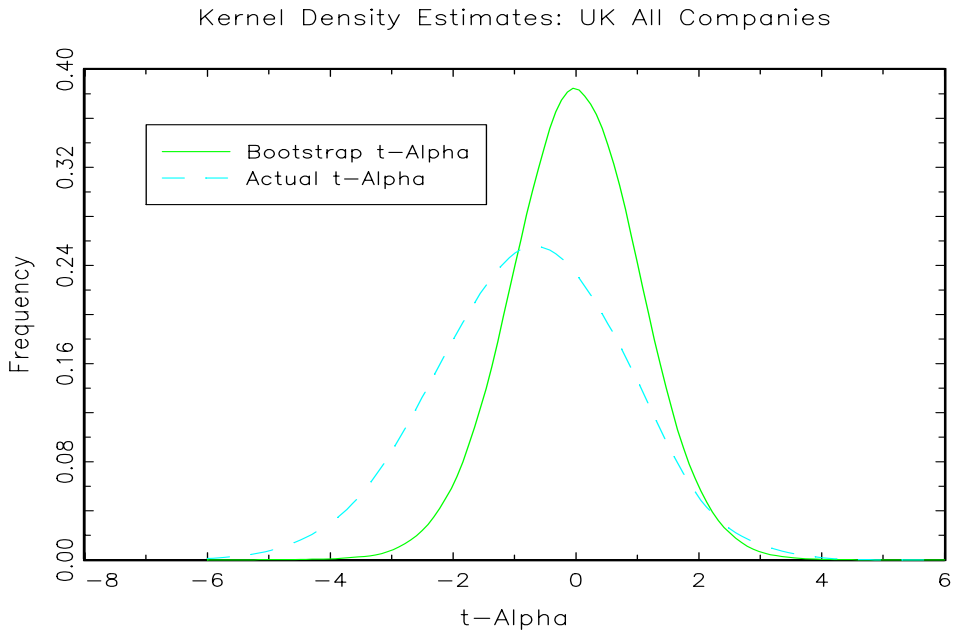
Figure 7. Kernel Density Estimates of the Actual and Bootstrap distribution of t-alphas - by Investment Style

Figure 7 shows Kernel densities of 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. t-statistics are Newey-West adjusted and funds with a minimum of 36 observations are used. The plots are generated using a Gaussian Kernel.

Equity Income



UK All Companies



Smaller Companies

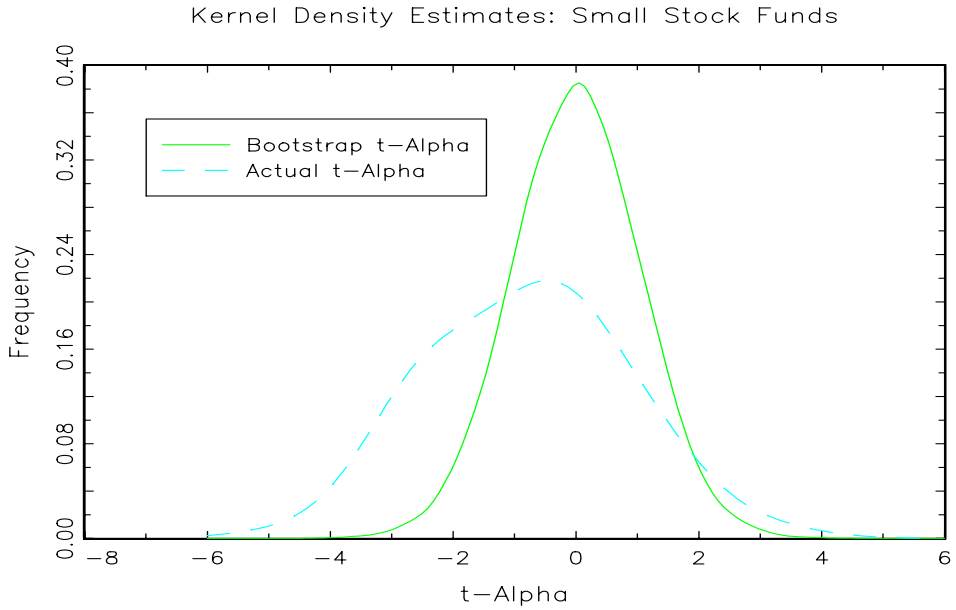
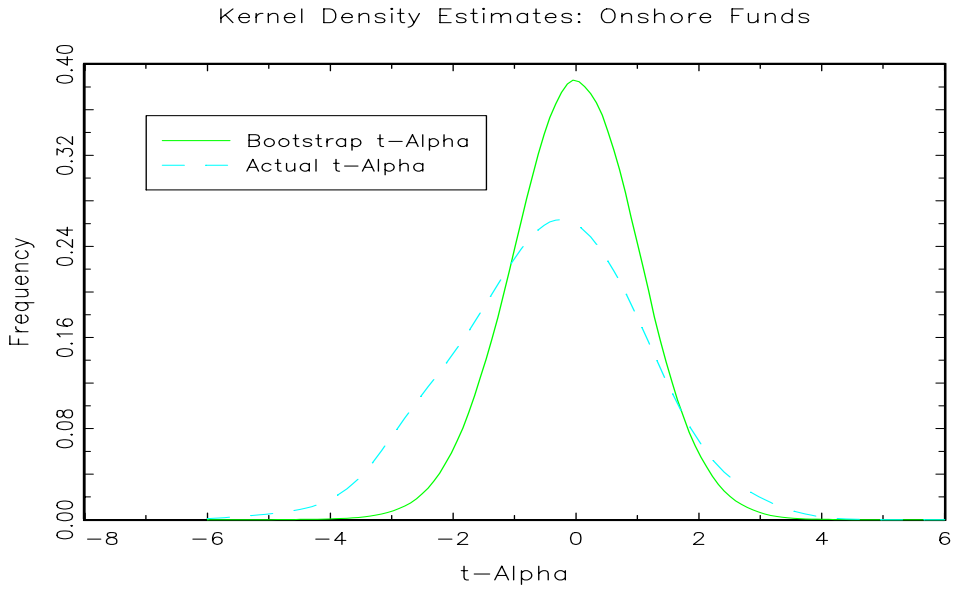


Figure 8. Kernel Density Estimates of the Actual and Bootstrap Distribution of t-alphas - by Location

Figure 8 shows Kernel densities of 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 36 observations are used. The plots are generated using a Gaussian Kernel.

Onshore Funds



Offshore Funds

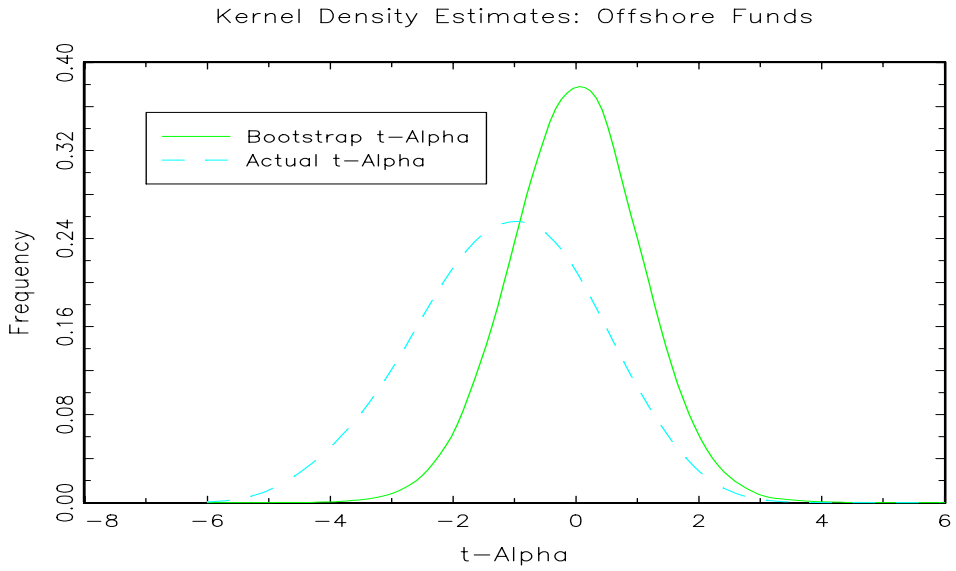
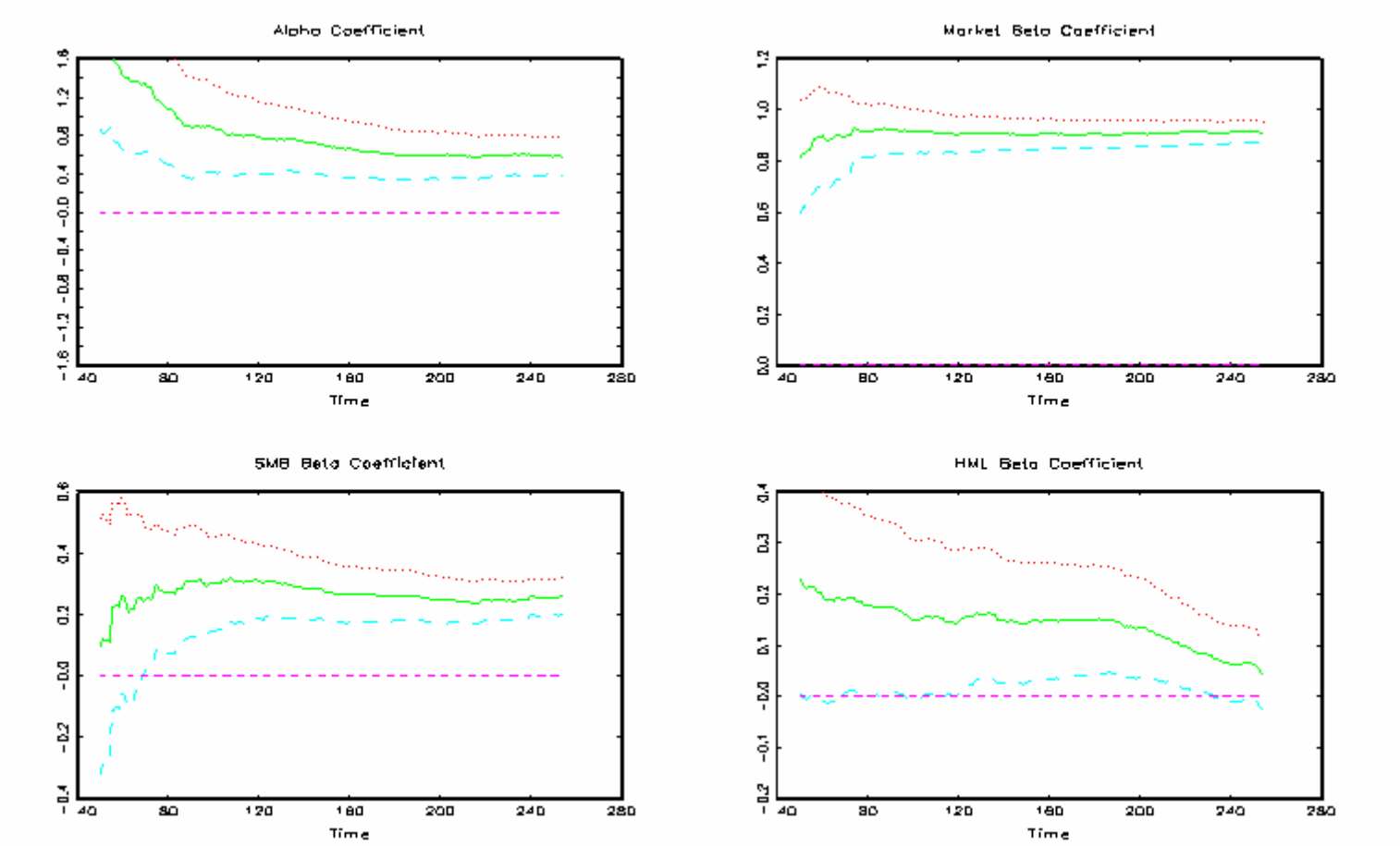
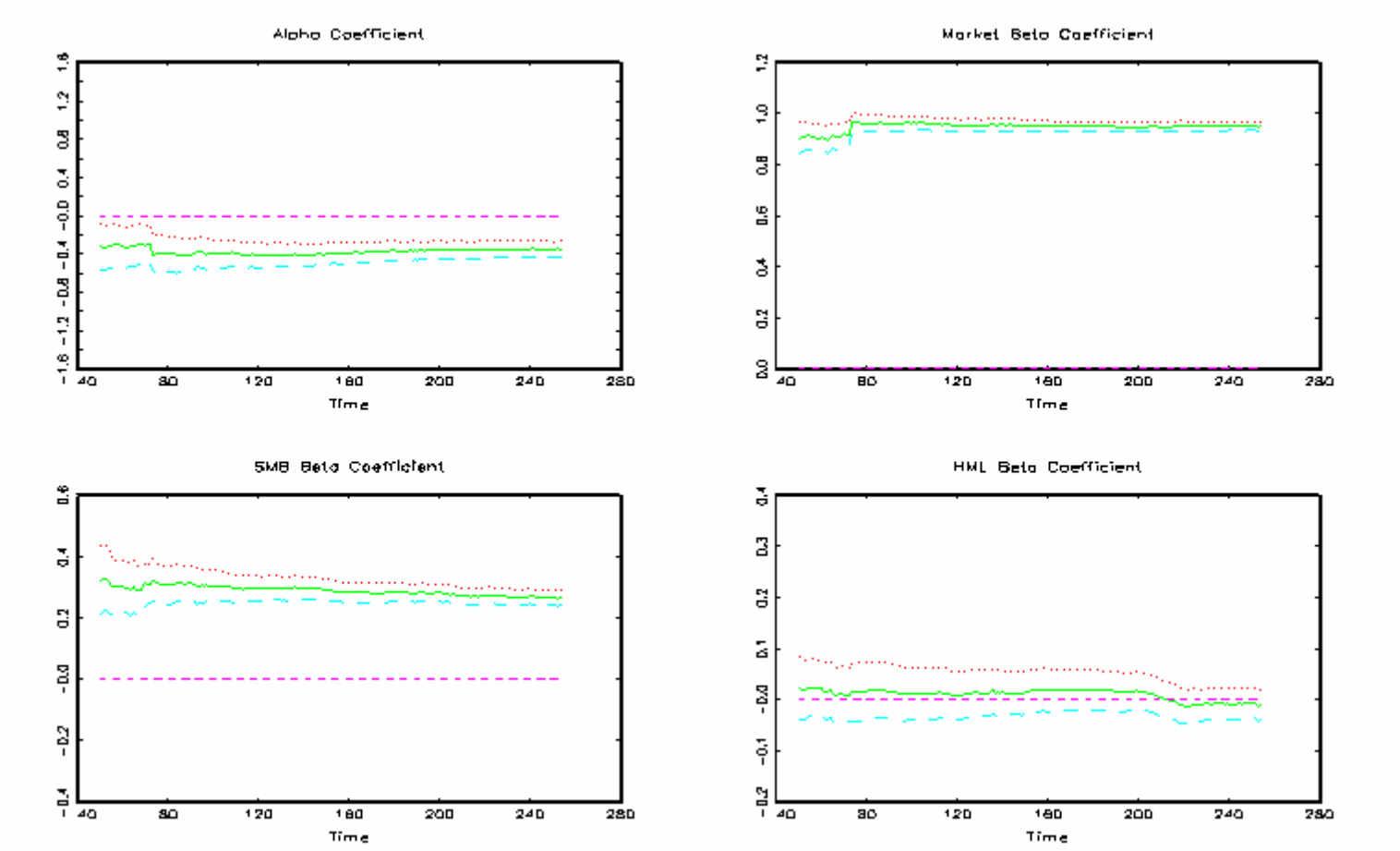


Figure 9. Recursive Estimates of Parameters (Portfolio of 'Best 12 Funds') : October 1981 – December 2002



Note : The 'best 12 funds' are selected by their bootstrapped t-alphas (10% level of significance), using the whole data set.

Figure 10. Recursive Estimates of Parameters (Portfolio of 'Worst 50 Funds') : October 1981 – December 2002



Note : The '50 worst funds' are selected by their bootstrapped t-alpha (using the whole data set).