

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, we use a cross-section 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 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, Ales Cerny, Stuart Hyde, Turalay Kenc, Aneel Keswani, Ian Marsh, Todd Milbourn, David Miles, Miles Gulnar Muradoglu, Andrew Sykes (FSA), Dylan Thomas, Giovanni Urga, Guofu Zhou 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; European Financial Management Association, Madrid June 2006; Olin Business School, St Louis October 2006; Barclays Capital (London), October 2006; AFFI conference, Paris, December 2006.

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), although it is worth pointing out that the power properties of standard tests of abnormal performance are quite low, even for relatively high levels of abnormal performance (e.g. 3% p.a., Kothari and Warner 2001).

A second major issue is whether abnormal performance can be identified ex-ante and for how long it persists. 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 Zechauer 1993, Brown and Goetzmann 1995, Carhart 1997, Wermers 2003, Blake and Morey 2000, Bollen and Busse 2004, 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. US 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, Wermers 2003). 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, Cuthbertson et al 2006).

This study examines the ex-post 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 ex-post performance of UK mutual funds (unit trusts). Note that we focus on the ex-post performance of *individual* funds

and not on ex-ante sorting rules that may yield evidence of performance persistence (in portfolios of funds).

A key contribution of the paper is for the first time, to assess the ex-post performance of *individual* UK funds and to demonstrate that standard tests may give misleading inferences, particularly in the tails of the performance distribution. 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 which has not been applied to UK data (and was first applied to US mutual funds by Kosowski, Timmermann, White and Wermers 2007). We use ‘alpha’ α and the t-statistic of alpha t_α , as our measures of risk adjusted performance. However, we do not assume as many earlier studies do, that a fund’s idiosyncratic risk has a known parametric distribution - we separate ‘skill’ from ‘luck’ even when the distribution of idiosyncratic risk across funds is highly non-normal. Our cross-section bootstrap allows us to obtain a performance distribution for all funds some of which are in the extreme tails of the cross-section distribution – precisely the funds that investors are likely to be most interested in (i.e. extreme ‘winners’ or ‘losers’) and for which conventional test procedures may be misleading.

In fact, we 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 examine a wide range of alternative models, control for survivor bias by including 236 ‘nonsurviving’ funds and perform a number of bootstrap techniques (to account for serial correlation, heteroscedasticity and possible contemporaneous cross-section correlation across the idiosyncratic risk of the funds).

We now anticipate some of our key findings. In contrast to earlier UK studies the bootstrap procedure indicates there is strong evidence in support of stock picking ability but only for 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, 12 funds exhibit levels of performance which cannot be attributable to ‘luck’ at 10% significance level¹. 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

¹ For the moment we ignore the issue of the appropriate significance level when in a multiple testing framework (Storey 2002).

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 (Barberis and Thaler 2003).

When examining different fund 'styles', we find outperformance for 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 skill, whereas the positive alphas for the best 'offshore funds' are due to luck. In the left tails of these distributions, we find 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 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.

In the rest of the paper we proceed as follows. Section I describes the data used in the study. In section II we discuss performance measurement models applied to mutual fund returns. Section III details the bootstrap methodology. In section IV we evaluate the performance measurement models and select a subset of 'best models' to which we apply the bootstrap procedure. Section V examines the results of the bootstrap analysis and section VI concludes.

I. 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 have removed '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, which exist for some or all of the complete data period.

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) and we can identify onshore funds (731) which are managed in the UK and offshore funds (204) which 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. Hence, we follow the usual convention in using net returns (bid-price to bid-price, with gross 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 index². 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 index³. The factor mimicking portfolio's momentum behavior, MOM, has been constructed using the constituents of the London Share Price Data Base, (total return) index⁴. 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 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.

³ These indices are constructed by Morgan Stanley who ranks all the stocks in their UK national index by their book-to-market ratio. Starting with the highest book-to-market ratio stocks, these are attributed to the value index until 50% of the market capitalization of the national index is reached. The remaining stocks are attributed to the growth index. The MSCI national indices have a market coverage of at least 60% (more recently this has been increased to 85%). Total return indices are used for the construction of the HML variable.

⁴ For each month, the equally weighted average returns of stocks with the highest and lowest 30% returns, over the previous eleven 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.

II. 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.

A. Unconditional Models

These have factor loadings that are time invariant. Carhart’s (1997) performance measure is the alpha estimate from a four-factor model:

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

where $r_{i,t} = (R_{i,t} - R_{f,t})$, $R_{i,t}$ = return on fund-i in period t , $R_{f,t}$ = risk free rate, $r_{m,t} = (R_{m,t} - R_{f,t})$ is the excess return on the market portfolio while SMB_t , HML_t and MOM_t are factor mimicking portfolios for size, book-to-market value and momentum effects, respectively. $\alpha_i = 0$ indicates zero abnormal performance. 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 Jensens (1968) 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 Siquefield 2000, Tonks 2004).

B. Conditional-Beta Models

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

$$(2) \quad r_{i,t+1} = \alpha_i + b_{0i}(r_{b,t+1}) + B_i'(z_t * r_{b,t+1}) + \varepsilon_{i,t+1}$$

where $r_{b,t+1}$ = the excess return on a benchmark portfolio (i.e. market portfolio in this case).

C. 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:

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

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.

D. 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 $\beta_{it} = \theta_i + \gamma_{im}[r_{m,t}]$ where $\gamma_{im} > 0$ is the unconditional measure of market timing ability. Alternatively, the Merton and Henriksson (1981) model of market timing uses $\beta_{it} = \theta_i + \gamma_{im} \max\{0, r_{m,t}\}$. 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.

III. Bootstrap Methodology

Previous studies of the ex-post performance of UK mutual funds all use ‘conventional’ statistical measures, and generally find (using a three or four factor model) that there is little or no positive abnormal performance in (portfolios of) the ‘best’ funds, whereas (portfolios of) the ‘worst’ funds have statistically significant negative risk adjusted performance (see *inter alia*, Blake and Timmermann 1998, Quigley and Sinquefeld 2000, Fletcher and Forbes 2002). Similarly, 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 ordered* funds, even when idiosyncratic risks are highly non-normal. Kosowski et al (2007) 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: $r_{i,t} = \hat{\alpha}_i + \hat{\beta}_i' X_t + e_{i,t}$ for $i = \{1, 2, \dots, n\}$ funds, where T_i = number of observations on fund- i , $r_{i,t}$ = excess return on fund- i , X_t = matrix of risk factors and $e_{i,t}$ are the residuals. 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 $\{\hat{\beta}_i, e_{i,t}\}$. 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 ($\alpha_i = 0$): $\tilde{r}_{i,t} = 0 + \hat{\beta}_i' X_t + \tilde{e}_{i,t}$. This is repeated for all funds. Next, using the simulated returns $\tilde{r}_{i,t}$, the performance model is estimated and the resulting estimate of alpha $\tilde{\alpha}_i^{(1)}$ for each fund is obtained. The $\tilde{\alpha}_i^{(1)}$ estimates for each of the n -funds represent sampling variation around a true value of zero (by construction) and are entirely due to ‘luck’. The $\tilde{\alpha}_i^{(1)}$ $\{i = 1, 2, \dots, n\}$ are then ordered from highest to lowest, $\tilde{\alpha}_{\max}^{(1)}$ to $\tilde{\alpha}_{\min}^{(1)}$. The above process is repeated $B = 1,000$ times for each of the n -funds which gives a separate ‘luck distribution’ for each of the ordered funds $f(\tilde{\alpha}_i)$ 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 $f(\tilde{\alpha}_{\max})$, we reject the null that its performance is due to luck (at 95% confidence) and infer that the fund has skill. This can be repeated for any other point in the performance distribution, right down to the ex-post worst performing fund in the data⁵. Note that the distribution $f(\tilde{\alpha}_{\max})$ 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. 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 ignores the other luck distributions encountered by all other funds – these other ‘luck distributions’ provide highly valuable and relevant information. Another key difference in our study 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 follow any distribution (which can be different for each fund). Hence the distribution under the null $f(\tilde{\alpha}_{\max})$, 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, respectively⁶.

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 $t_{\tilde{\alpha}_i}$,

⁵ For the top performing funds we test the null hypothesis $H_0: \alpha_i \leq 0$, $H_A: \alpha_i > 0$ and for the bottom performing funds $H_0: \alpha_i \geq 0$, $H_A: \alpha_i < 0$.

⁶There 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 (2005), co-skewness of individual constituent non-normal security returns may not be diversified away in a fund. Also, funds may hold derivatives to hedge return outcomes and this may result in a non-normal return distribution. 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.

which is important for inference in the extreme tails (Kosowski et al 2007, Hall 1992)⁷. 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 $T_i \geq 36$ months.

IV. Model Selection

Table I provides a summary of the number of funds, as well as the three and four factor alphas (for equally weighted portfolios), over consecutive 3 year non-overlapping periods. These performance statistics are presented for those funds with $T_i \geq 36$ observations and also for all the funds in our database. Panel A reports results for all fund styles while panels B, C and D report results for three different investment styles. Restricting the analysis to funds with $T_i \geq 36$ reduces the number of funds in most years by between 10-25% but this does not alter the 3-factor and 4-factor alphas by more than 25 basis points per annum, for the whole set of funds (Panel A). For income funds and all-UK companies funds the differences in alphas (Panels B and C) for all funds and those with $T_i \geq 36$ are again around 25 basis points but for small company funds these differences are a little higher but still small. Thus our attempt to minimize survivorship bias by only using funds with $T_i \geq 36$ should not distort our results from the bootstrap (but below we also report variants with $T_i \geq 18$ and 24). It is also worth noting that (for funds with $T_i \geq 36$) the three and four factor alphas do not differ greatly and this is also found to be the case (see below) when we investigate the alphas for individual funds.

[Table I here]

In this section, the equilibrium models described in section II 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 $T_{i,\min} = 36$. 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, using the Schwartz Information Criterion (SIC) and these

⁷ t_{α_i} has better sampling properties than α_i - the obvious reason being that the former 'corrects for' high risk-taking funds (i.e. σ_{ϵ_i} large), which are likely to be in the tails. If different funds have different distributions of idiosyncratic risk (e.g. different skewness and kurtosis) then we cannot say a priori what the distribution of $f(t_{\alpha_i})$ will be – this is why we need the cross-section bootstrap.

results are reported in table II. (We examined many variants within the three classes of model and in none of our models are the Treynor-Mazuy (1966) and Merton-Henriksson (1981) market timing variables significant - these results are available on request).

[Table II here]

In the best three models (bottom half of table II), the cross-sectional average alpha takes on a small and statistically insignificant negative value (consistent with Blake and Timmermann 1998). 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 II) 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 II), 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.

V. 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 II, we report that around 64% of mutual funds reject normality in their regression residuals and (in results not reported) 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). This is evident in figures 1a and 1b which show the bootstrap histograms of $t_{\hat{\alpha}_i}$ at selected points 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.

[Figures 1a and 1b here]

Table III 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 $t_{\hat{\alpha}_i}$ under the null of no outperformance.

[Table III here]

For example, using the *unconditional* model the 'max' fund (Table III, 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 1a, 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 $t_{\hat{\alpha}}$ of the best fund indicates 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 III, panel A.

A straight ‘count’ from our complete set of bootstrap results show that of the top 20 ranked funds, 12 achieve outperformance (each at a 10% significance level) while 7 funds outperform (each at a 5% significance level)⁸. 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 III, panel A and figure 1a).

In the left tail of the distribution, (i.e. the left side of Panel A, table III), the lowest ranked fund has $t_{\hat{\alpha}} = -5.358$ with a bootstrap p-value of 0.009 - this fund has produced ‘truly’ inferior performance. This can be seen in the upper left panel of figure 1b, 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 III (and figure 1b), that all funds in the left tail (up to the ‘min 40%’ point) have ‘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’. Figure 2 shows 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. 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’. For example, based on the unconditional (FF) model we would expect 10 funds

⁸ Statements like this are frequently made in the literature but must be interpreted with caution. The bootstrap correctly computes the p-value of the *individual* ordered funds. However, when we count the number of significant funds (each at a particular significance level such as 5%) we implicitly undertake a multiple-testing problem and some of these ‘significant’ funds may be significant due to pure chance. Put another way, if the significance level of each individual fund is set at 5%, the probability of finding at least one fund out of M -significant funds is greater than 5%. Hence, statements such as this in the text, provide an upper bound for the *number* of truly successful funds.

to achieve $\hat{\alpha} \geq 0.5\%$ p.m. (6% p.a.) based on random chance alone, whereas 19 funds exhibit this level of performance (or higher). However, $\hat{\alpha} \geq 0.1\%$ 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 outperformance *just within* the extreme right tail, than nearer the centre of the performance distribution.

[Figure 2 here]

Panels B and C of Table III 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 III) are broadly consistent with those for the unconditional model (Panel A). Skill is found for around 8% 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.

There is also some sensitivity in the bootstrap results to potential survivorship bias. For example, when we repeat the analysis of table III for the unconditional 3F model but include funds with a minimum number of observations of either 24 or 18 months, then the number of funds with skill (at a 10% significance level) are 16 and 22 respectively, which constitute 2.2% and 3.0% respectively of all funds included. This compares with 12 winners (i.e. 1.8% of all funds) when $T_i \geq 36$. Results for the conditional beta model show that the proportion of skilled funds for $T_i \geq 36, 24$ or 18, are 7.8%, 8.7% and 10.4% respectively, while for the conditional alpha-beta model the equivalent figures are 4.1%, 3.8% and 3.7% of funds (at a 10% significance level). For the left tail of the performance distribution, results with $T_i \geq 24$ or 18 clearly show that funds up to the 40th percentile have bad skill rather than bad luck – results that are qualitatively similar to those reported for $T_i \geq 36$. So, as one might expect, there is some variation in the bootstrap results when more funds are included but the broad qualitative results in table III remain unchanged – for both tails 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 III, while addition of the momentum variable produces slightly more winners (around 5%) than the unconditional 3 factor model (of

table III, 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 (2007) and Barras, Scaillet and Wermers (2005), who find strong evidence of stock picking ability among the very top performing of US funds and poor performance for many funds in the left tail of the performance distribution⁹.

Above we applied the bootstrap across all funds using each of our 3 ‘best models’. However, recall from Table II 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’.

A. 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 III), 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. Above we noted that the proportion of winner and loser funds identified in the bootstrap is qualitatively unchanged as we reduce the minimum data length from 36 to 24 or 18 observations - so survivorship bias does not appear to be severe.

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. 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.

⁹ Kosowski et al (2005) use a cross-section bootstrap and based on a straight count for ‘all funds’, they find between 5-10% of funds (depending on the model chosen) have genuine skill (each at a 5% significance level or better). Barras et al (2005) do not require a multivariate bootstrap and measure luck using the ‘false discovery rate’ FDR - that is, the proportion of lucky funds among funds with significant (individual) alphas. Using the FDR they find only about 2% of funds have genuine skill (at a 5% significance level). Both studies find that genuinely poor performing funds are spread throughout much of the left tail. Note however, that the Kosowski et al (2005) and Barras et al (2005) definitions of luck are very different. The former provides evidence on the performance of individual *ordered* funds (based on correctly estimated p-values) - as in this study. Barras et al (2005) provide a correct method for calculating the *proportion* of truly successful funds from a set of ‘significant’ funds – however, it is not possible to identify which of the individual funds are truly successful (assuming we rule out the extreme case of an FDR of zero).

B. Performance and Investment Styles

Having found some 'good skill' and lots of 'bad skill' when analyzing all UK mutual funds, it is interesting to see whether skill (or lack of it) is 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 IV here]

Table IV, Panels A, B and C re-estimate the performance statistics of table III, for the three investment styles. Looking at the left side of all three Panels in Table IV ('t-alpha', 'p-tstat') it is clear that '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 IV, 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 IV, Panels B and C, 't-alpha', 'p-tstat'), there are hardly any funds which have 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')¹⁰. Our findings for the UK of 'skill' mainly among some top performing UK *income* funds are in contrast to results in Kosowski et al (2007) and Barras et al (2005) for US mutual funds, who find that it is the top performing *growth* funds that have skill. (But note that these US studies do not have a 'small companies' style classification)¹¹.

¹⁰The above results are consistent with those in table III, 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 III income funds are proportionately more representative of skill, than the other fund styles.

¹¹In results not reported the Kernel density estimates for the three investment styles give similar qualitative results to those in table IV. 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 of the best funds does not exceed random sampling variation. Also 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. 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

C. 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. Differential performance may arise due to possible information asymmetries between offshore versus onshore operations or simply differential skill given identical information.

[Table V here]

The bootstrap results in table V ('t-alpha', 'p-tstat') are clear cut. Out of 553 onshore funds almost all of the top 20 possess 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'¹². Skill for the top-20 onshore funds but not the offshore funds is in small part due to differences in fees since offshore funds ('all company') have expenses which are around 0.5% p.a. higher than onshore funds (at end-2002) and our returns data is net of management expenses. Also, onshore funds may have an informational advantage resulting in more successful asset allocation (as documented in some US studies such as Coval and Moskowitz 1999, Nanda et al 2004 and Kacperczyk 2005).

Very wealthy investors may hold the top offshore funds even though their pre-tax (risk adjusted) return is small (and statistically insignificant) because any net (absolute) returns may involve lower (or near zero) personal taxes compared to onshore funds. But given that effective capital gains tax payments for most UK investors in onshore funds are probably quite small¹³, it is unlikely that they would rationally hold top performing offshore funds, when they would be better off in the top performing onshore funds. A similar anomaly arises for most funds in the left tail

a given level of performance from among the actual funds, against the frequency of this same level of performance in the *entire* bootstrap distribution. Also, see Barras et al (2005) on the use of the false discovery ratio, in this context. 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.

¹² These results are also consistent with those of table III, where all of the 12 'skilled' funds (FF unconditional model, 10% significance level) can be identified as onshore funds.

¹³ Capital gains taxes on individual UK investors from onshore funds can be completely avoided if funds are held as part of a PEP/ISA, which began in the mid-1990s (with an annual limit of around £7,500 p.a. and where dividends can be re-invested). Also, UK capital gains tax is only paid on gains in excess of £8,200 (2004/5) in any one year, on any onshore investments held outside PEP/ISA's. In practice therefore, for many individual UK investors the tax advantages of offshore as opposed to onshore funds is limited. Of course the relative attraction of offshore UK funds to other nationalities depends on their own domestic tax rates and reciprocal tax arrangements with offshore centres. For example, Jersey (like Switzerland) is not a member of the European Union and does not necessarily have to comply with its information-sharing rules – although it has agreed to the EU Savings Tax Directive and a withholding tax became operative from 2005. There is a double taxation agreement with the UK but not generally with other countries. Perhaps a major attraction of offshore funds is secrecy. Assets held under 'beneficial ownership' in Jersey by non-residents must be disclosed to the Jersey authorities but the information is not kept on public record and cannot be disclosed externally except by order of the Royal Court. Hence 'secrecy' may provide an implicit return to some offshore investors – as demonstrated for Switzerland by English and Shahin (1994) and Delaloye, Habib and Ziegler (2005).

(whether onshore or offshore – table III) - since these funds have ‘bad skill’, yet stay in existence. 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 (2007) also find strong evidence of 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.

As with other authors we can only conjecture that the survival of poorly performing funds may be due to incomplete information, inertia, disposition/endowment effects or some element of irrationality on the part of investors, as documented in the behavioural finance and mutual fund literature¹⁴. Our results also imply that in the context of the Berk and Green (2004) model any move towards a competitive equilibrium among poorly performing funds, appears to be relatively slow.

D. Extensions of the Bootstrap

The ‘baseline’ bootstrap procedure described in section IV 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.

E. 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), *after* any ‘winners’ or ‘losers’ have been identified by the bootstrap

¹⁴

An excellent survey of the behavioural finance literature is Barberis and Thaler (2003) and for recent examples applied to mutual funds see for example Cooper, Gullen and Rau (2005) on fund flows and cosmetic name changes, Hong, Kubik and Stein (2005) for word-of-mouth effects on fund’s stock selection and Elton, Gruber and Busse (2004) on the relative performance within index funds.

methodology. The alternative is that the parameters are statistically significant in the bootstrap but are not constant over time. Such evidence would complement our bootstrap results which use 't-alpha' based on the *whole sample*, as the performance criterion. Investors want to be reasonably sure that any (ex-post), 'alpha-performance' is not sample specific. Note that this provides a 'descriptive analysis' rather than a 'predictive analysis', as we are not looking at portfolio performance from an *ex-ante* viewpoint as in studies of performance persistence.

The problem in doing even a descriptive analysis of parameter constancy is that some funds are born and some die (or are merged), so that the *number* of winner/loser funds varies as we move through time. We use recursive estimates as a useful way to encapsulate the broad thrust of what is happening to the parameters of winner or loser portfolios¹⁵. Recursive OLS (with GMM correction for standard errors) and the Kalman filter model are used with our unconditional Fama-French 3 factor model. 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: $\alpha_t = \alpha_{t-1} + v_t$, $\beta_{k,t} = \beta_{k,t-1} + v_{k,t}$ ($k = 1, 2, 3$). This specification allows the $\alpha_t, \beta_{k,t}$ to change considerably from period-to-period depending on the variances of the error terms σ_v , which we assume are contemporaneously uncorrelated. (Recursive OLS is a special case of the Kalman filter where $\sigma_v = 0$ - see Hamilton 1994).

For an equally weighted portfolio of the best 12 funds (as identified in the bootstrap), the recursive OLS estimates over the period June 1981 to December 2002 (with GMM corrected standard errors) are shown in figure 3¹⁶. 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 constant estimate of alpha of around 0.58 (6.96% p.a.) over the whole recursive period to end December 2002.

[Figure 3 and 4 here]

At the bottom end of the performance distribution (figure 4) 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

¹⁵ Recently, recursive estimation has been interpreted by Mamaysky, Spiegel and Zhang (2004) as an unobservable variables approach, which is complementary to the use of conditional models in that both approaches hope to mimic dynamic trading strategies.

¹⁶ In the recursive plots the number of funds in the equally weighted portfolio varies. For the 'best funds' identified by the bootstrap the minimum number in the recursive plots is 4 and the maximum number is 12, while for portfolio of 'worst funds' the minimum number of funds in the early years of the sample is 8 and the maximum rises to 50.

reasonably constant, while that on HML is again statistically insignificant. This general qualitative pattern is repeated when we include all of the worst funds (as indicated by our bootstrap procedure) up to the 40th percentile in our equally weighted portfolio.

[Table VI 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 VI we present the final state vectors (at time T), their p-values and the standard error (σ_v) of the time varying parameters (see Hamilton 1994). A portfolio of the best 12 funds (which *all* have actual t-alphas significant at a 10% critical value as indicated by the bootstrap) has a final state vector $\alpha_T = 0.446$ (p-value = 0.0083). The relatively low standard error $\sigma_v = 0.0279$ of the time varying alpha confirms that our ranking on bootstrap t-alphas does provide a portfolio of 'top' funds that exhibits constant alpha performance over the whole period of around 5.3% p.a. 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 ($\sigma_v = 0.0003$) and the final state vector is $\alpha_T = -0.337$ (around -4.0% p.a.) with p-value < 0.0001 – similar to the recursive OLS results.

VI. Conclusion

Using a comprehensive data set for UK equity mutual funds, April 1975 – December 2002, we use a novel cross-section bootstrap methodology to distinguish between 'skill' and 'luck'. Depending on the particular model chosen, we find 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 broadly consistent with recent US empirical evidence (Kosowski et al 2007, Barras et al 2005). 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 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 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 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 results strongly reject the view that most poor performing funds are merely unlucky and most of these funds demonstrate 'bad skill' - which is broadly consistent with results for US funds (Kosowski et al 2007, Barras 2005). However, this result is not consistent with the competitive model of Berk and Green (2004), since 'bad skill' should lead to large cash outflows 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 an element of irrationality.

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 a constant positive alpha for a portfolio of a few top funds and a constant negative alpha for a portfolio of many poorly performing funds. However, it would also be interesting to investigate the power of the cross-section bootstrap in detecting performance persistence in an ex-ante context, particularly in 'small' fund-of-fund winner/loser portfolios, where there is likely to be severe non-normality in idiosyncratic risk. Given the number of possible variants of potential sorting rules, rebalancing and holding periods, this remains a task for future research.

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. 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., S. Bhattacharya, Stephen A. Ross, and P. Pfleiderer, 1986, On Timing and Selectivity, *Journal of Finance* 41, 715-730.
- Allen, D.E., and M.L. Tan, 1999, A Test of the Persistence in the Performance of UK Managed Funds, *Journal of Business Finance and Accounting* 25, 559-593.
- Baks, K.P., A. Metrick, and J. Wachter, 2001, Should Investors Avoid All Actively Managed Funds? A Study in Bayesian Performance Evaluation, *Journal of Finance* 56, 45-86.
- Ball, R., and S.P. Kothari, 1989, Nonstationary Expected Returns: Implications for Tests of Market Efficiency and Serial Correlations in Returns, *Journal of Financial Economics* 25, 51-74.
- Barberis, Nicholas, and R.H. Thaler, 2003, A Survey of Behavioral Finance, in G.M. Constantinidis, M. Harris and R. Stulz, ed.: *Handbook of the Economics of Finance*, Elsevier Science B.V.
- Barras, Laurent, Olivier Scaillet, and Russ Wermers (2005) False Discoveries in Mutual Fund Performance: Measuring Luck in Estimated Alphas, FAME Research Paper No.163, University of Geneva.
- Berk, Jonathan B., and Richard C. Green, 2004, Mutual Fund Flows and Performance in Rational Markets, *Journal of Political Economy* 112, 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, Christopher A., and M. Morey, 2000, Morningstar Ratings and Mutual Fund Performance, *Journal of Financial and Quantitative Analysis* 35, 451-483.
- Blake, David, and Allan Timmermann, 1998, Mutual Fund Performance : Evidence from the UK, *European Finance Review* 2, 57-77.
- Blake, David, and Allan Timmermann, 2003, Performance Persistence in Mutual Funds : An Independent Assessment of the Studies Prepared by Charles River Associates for the Investment Management Association, *Financial Services Authority* (also available at www.pensions-institute.org/reports)
- Blake, David, Bruce N. Lehmann, and Allan Timmermann, 1999, Performance Measurement Using Multi-Asset Portfolio Data: A Study of UK Pension Funds 1986-94, Pensions Institute, London.
- Bollen, Nicolas P.B., and Jeffrey A. Busse, 2004, Short-Term Persistence in Mutual Fund Performance, *Review of Financial Studies*, 18(2),569-597
- Brown, Stephen J., and William N. Goetzmann, 1995, Performance Persistence, *Journal of Finance* 50, 679-698.
- Brown, Stephen J., William N. Goetzmann, Roger G. Ibbotson, and Stephen A. Ross, 1992, Survivorship Bias in Performance Studies, *Review of Financial Studies* 5, 553-580.
- Carhart, Mark M, J. Carpenter, A. Lynch, and D. Musto, 2002, Mutual Fund Survivorship,

The Review of Financial Studies 15, 1439-1463.

- Chen, Hsiu-Lang, Narasimhan Jegadeesh, and Russ Wermers, 2000, The Value of Active Mutual Fund Management : An Examination of the Stockholdings and Trades of Fund Managers, *Journal of Financial and Quantitative Analysis* 35, 343-368.
- Chevalier, J., and G. Ellison, 1999, Are Some Mutual Fund Managers Better than Others? Cross-Sectional Patterns in Behavior and Performance, *Journal of Finance* 54, 875-899.
- Christopherson, Jon A., Wayne E. Ferson, and Debra A. Glassman, 1998, Conditioning Manager Alphas on Economic Information: Another Look at the Persistence of Performance, *Review of Financial Studies* 11, 111-142
- Cooper, Michael J., Huseyin Gulen and P. Raghavendra Rau, 2005, Changing Names With Style: Mutual Fund Name Changes and Their Effects on Fund Flows, *Journal of Finance*, 60, 2825-2858.
- Coval, Joshua D., and Tobias J. Moskowitz, 1999, Home Bias at Home: Local Equity Preference in Domestic Portfolios, *Journal of Finance* 54, 2045-2074.
- Cuthbertson, Keith, Dirk Nitzsche and Niall O'Sullivan, 2006, Mutual Funds: Persistent Winners or Losers?, *Cass Business School, Discussion Paper*.
- Daniel, Kent M., Mark Grinblatt, Sheridan Titman, and Russ Wermers, 1997, Measuring Mutual Fund Performance With Characteristic Based Benchmarks, *Journal of Finance* 52, 1035-1058.
- Delaloye, Francois-Xavier, Michel Habib and Alexandre Ziegler, 2005, Negotiating Over Banking Secrecy: the Case of Switzerland and the European Union, Swiss Banking Institute, University of Zurich, Working Paper.
- Efron, B., and R.J. Tibshirani, 1993. *An Introduction to the Bootstrap*, Monographs on Statistics and Applied Probability (Chapman and Hall, New York).
- Elton, Edwin J., Martin J Gruber and Jeffrey A. Buse, 2004, Are Investors Rational? Choices Among Index Funds, *Journal of Finance*, 59, 261-288.
- English, Mary and Wassim Shahin, 1994, Investigating the Interest Rate Impact of Changing Secret Bank Deposit Laws: Switzerland, *Journal of Banking and Finance*, 18, 461-475.
- Fama, Eugene F., and Kenneth R. French, 1993, Common Risk Factors in the Returns on Stocks and Bonds, *Journal of Financial Economics* 33, 3-56.
- Ferson, Wayne E., and Rudi W. Schadt, 1996, Measuring Fund Strategy and Performance in Changing Economic Conditions, *Journal of Finance* 51, 425-62.
- Fletcher, Jonathan, 1995, An Examination of the Selectivity and Market Timing Performance of UK Unit Trusts, *Journal of Business Finance and Accounting* 22, 143-156.
- Fletcher, Jonathan, and David Forbes, 2002, An Exploration of the Persistence of UK Unit Trusts Performance, *Journal of Empirical Finance* 9, 475-493.

- Foster, G., 1986. *Financial Statement Analysis* (Prentice Hall, Englewood Cliffs).
- Goetzmann, William N., Z. Ivkovic, and G. Rouwenhorst, 2001, Day Trading International Mutual Funds: Evidence and Policy Solutions, *Journal of Financial and Quantitative Analysis* 36, 287-310.
- Grinblatt, Mark, and Sheridan Titman, 1992, 'The Persistence of Mutual Fund Performance, *Journal of Finance* 47, 1977-1984.
- Grinblatt, Mark, Sheridan Titman, and Russ Wermers, 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, Hamburg).
- Hamilton, J.D., 1994. *Time Series Analysis* (Princeton University Press, Princeton, N.J.).
- Hendricks, Darryll, Jayendu Patel, and Richard Zeckhauser, 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.
- Hong, Harrison, Jeffrey D. Kubik and Jeremy C. Stein 2005, Thy Neighbor's Portfolio: Word-of-Mouth Effects in the Holdings and Trades of Money Managers, *Journal of Finance* 60, 2801-2824.
- Jensen, Michael C., 1968, The Performance of Mutual Funds in the Period 1945-1964, *Journal of Finance* 23, 389-416.
- Jiang, Wei, 2003, A Non Parametric Test of Market Timing, *Journal of Empirical Finance* 10, 399-425.
- Kacperczyk, Marcin, Clemens Sialm and Lu Zheng, 2005, On the Industry Concentration of Actively Managed Mutual Funds, *Journal of Finance*, LX(4), 1983-2011.
- Kosowski, Robert, Allan Timmermann, Hal White, and Russ Wermers, 2007, Can Mutual Fund "Stars" Really Pick Stocks? New Evidence from a Bootstrap Analysis, *forthcoming Journal of Finance*.
- Kothari, S.P and Jerold Warner, 2001, Evaluating Mutual Fund Performance, *Journal of Finance*, LVI (5), 1985-2010
- Lakonishok, Josef A., Andrei Shleifer, and Robert W. Vishny, 1992, The Structure and Performance of the Money Management Industry, *Brookings Papers on Economic Activity*, 339-391.
- Leger, L., 1997, UK Investment Trusts : Performance, Timing and Selectivity, *Applied Economics Letters* 4, 207-210.
- Mamaysky, Harry, Matthew Spiegel, and Hong Zhang, 2004, Improved Forecasting of Mutual Fund Alphas and Betas, Yale School of Management, ICF Working Paper 04-23.

- Mandelker, G.N., and S.G. Rhee, 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, Robert C., and R. Henriksson, 1981, On Market Timing and Investment Performance : Statistical Procedures for Evaluating Forecasting Skills, *Journal of Business* 54, 513-533
- Nanda, Vicram, Zhi Wang, and Lu Zheng, 2004, Family Values and the Star Phenomenon, *Review of Financial Studies*, 17(3), 667-698.
- Newey, Whitney D., and Kenneth D. West, 1987, A Simple, Positive Semi-Definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* 55, 703-708.
- Pastor, Lubos, and Robert F. Stambaugh, 2002, Mutual Fund Performance and Seemingly Unrelated Assets, *Journal of Financial Economics* 63, 315-350.
- Politis, D.N., and J.P. Romano, 1994, The Stationary Bootstrap, *Journal of the American Statistical Association* 89, 1303-1313.
- Quigley, Garrett, and Rex A. Siquefield, 2000, Performance of UK Equity Unit Trusts, *Journal of Asset Management* 1, 72-92
- Storey, John, D., 2002, A Direct Approach to False Discovery Rates, *Journal of the Royal Statistical Society*, 64,479-498.
- Teo, M., and Sung-Jun Woo, 2001, Persistence in Style-Adjusted Mutual Fund Returns, Manuscript, Harvard University.
- Thomas, A., and Ian Tonks, 2001, Equity Performance of Segregated Pension Funds in the UK, *Journal of Asset Management* 1, 321-343.
- Tonks, Ian, 2004, Performance Persistence of Pension Fund Managers, University of Exeter, Centre for Finance and Investment, Working Paper, forthcoming *Journal of Business* 2005.
- Treynor, Jack, and K. Mazuy, 1966, Can Mutual Funds Outguess the Market, *Harvard Business Review* 44, 66-86.
- Turner, Adair, 2004, Pensions : Challenges and Choices : The First Report of the Pensions Commission, The Pensions Commission, *The Stationary Office*, London
- Viceira, Luis M., 2001, Optimal Portfolio Choice for Long-horizon Investors with Nontradable Labor Income, *Journal of Finance* 56, 433-470.
- Wermers, Russ, 2000, Mutual Fund Performance : An Empirical Decomposition into Stock Picking Talent, Style, Transactions Costs, and Expenses, *Journal of Finance* 55, 1655-1703.
- Wermers, Russ, 2003, Is Money Really "Smart"? New Evidence on the Relation Between Mutual Fund Flows and Performance Persistence, Department of Finance, University of Maryland.

Table I : Summary Statistics of Mutual Funds

Table I shows the number of funds and their 3 and 4 factor alphas for an equally weighted portfolio of ‘All Funds’ and where funds have $T \geq 36$ observations and are estimated over 3 year non-overlapping periods. Estimation of the 4 factor models start in January 1982. The 3 factor model has the excess return on the market, the SMB and HML factors as explanatory variables, whereas the 4 factor model also includes the momentum variable.

Panel A : All Investment Objectives

	Number of Funds		3 Factor Model : alpha (equally weighted portfolio)		4 Factor Model : alpha (equally weighted portfolio)	
	All Funds	$T \geq 36$	All Funds	$T \geq 36$	All Funds	$T \geq 36$
Jan 1976 – Dec 1978	71	65	-0.6385	-0.6197	-	-
Jan 1979 – Dec 1981	163	72	-0.4648	-0.4485	-	-
Jan 1982 – Dec 1984	210	163	0.0318	-0.0011	0.0305	-0.0013
Jan 1985 – Dec 1987	311	211	0.0715	0.0442	0.1569	0.1140
Jan 1988 – Dec 1990	398	311	-0.2859	-0.3054	-0.2923	-0.3189
Jan 1991 – Dec 1993	463	398	-0.1288	-0.1389	-0.0929	-0.1011
Jan 1994 – Dec 1996	581	463	-0.0485	-0.0512	-0.0613	-0.0644
Jan 1997 – Dec 1999	694	385	-0.0258	-0.0367	-0.0342	-0.0453
Jan 2000 – Dec 2002	631	486	-0.0777	-0.0903	-0.0721	-0.0842

Panel B : Equity Income

	Number of Funds		3 Factor Model : alpha (equally weighted portfolio)		4 Factor Model : alpha (equally weighted portfolio)	
	All Funds	$T \geq 36$	All Funds	$T \geq 36$	All Funds	$T \geq 36$
Jan 1976 – Dec 1978	24	22	-0.7055	-0.6652	-	-
Jan 1979 – Dec 1981	51	25	-0.7324	-0.7072	-	-
Jan 1982 – Dec 1984	66	51	0.1817	0.1256	0.1885	0.1297
Jan 1985 – Dec 1987	92	66	0.2840	0.2343	0.3471	0.2982
Jan 1988 – Dec 1990	105	92	-0.2220	-0.2333	-0.2473	-0.2483
Jan 1991 – Dec 1993	110	105	-0.0672	-0.0678	-0.0092	-0.0109
Jan 1994 – Dec 1996	132	110	-0.1324	-0.1267	-0.1408	-0.1366
Jan 1997 – Dec 1999	144	80	-0.1285	-0.1266	-0.1151	-0.1132
Jan 2000 – Dec 2002	110	92	0.4064	0.4156	0.3607	0.3701

Panel C : UK All Companies

	Number of Funds		3 Factor Model : alpha (equally weighted portfolio)		4 Factor Model : alpha (equally weighted portfolio)	
	All Funds	$T \geq 36$	All Funds	$T \geq 36$	All Funds	$T \geq 36$
Jan 1976 – Dec 1978	42	38	-0.6114	-0.6039	-	-
Jan 1979 – Dec 1981	96	42	-0.3040	-0.2972	-	-
Jan 1982 – Dec 1984	124	96	0.0007	-0.0342	-0.0001	-0.0331
Jan 1985 – Dec 1987	178	125	0.0061	-0.0130	0.0794	0.0515
Jan 1988 – Dec 1990	238	178	-0.2500	-0.2679	-0.2910	-0.3262
Jan 1991 – Dec 1993	274	238	-0.1534	-0.1674	-0.1393	-0.1514
Jan 1994 – Dec 1996	351	274	-0.0593	-0.0523	-0.0698	-0.0647
Jan 1997 – Dec 1999	440	241	-0.0557	-0.00710	-0.0629	-0.0777
Jan 2000 – Dec 2002	430	318	-0.0974	-0.1147	-0.0952	-0.1116

Panel D : Smaller Companies

	Number of Funds		3 Factor Model : alpha (equally weighted portfolio)		4 Factor Model : alpha (equally weighted portfolio)	
	All Funds	$T \geq 36$	All Funds	$T \geq 36$	All Funds	$T \geq 36$
Jan 1976 – Dec 1978	5	5	-0.5399	-0.5399	-	-
Jan 1979 – Dec 1981	16	5	-0.4770	-0.4266	-	-
Jan 1982 – Dec 1984	20	16	-0.2247	-0.2062	-0.2529	-0.2282
Jan 1985 – Dec 1987	41	20	-0.0963	-0.2255	0.0899	-0.1030
Jan 1988 – Dec 1990	55	41	-0.5661	-0.6298	-0.3820	-0.4454
Jan 1991 – Dec 1993	79	55	-0.1176	-0.1516	-0.0227	-0.0556
Jan 1994 – Dec 1996	98	79	0.1041	0.0580	0.0784	0.0371
Jan 1997 – Dec 1999	110	64	0.2245	0.2050	0.1829	0.1613
Jan 2000 – Dec 2002	91	76	-0.5568	-0.6009	-0.4766	-0.5194

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

Table II shows results from the estimation of the performance models described in Section II 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 (25.196)	0.863 (21.193)	0.849 (21.068)
SMB_t	0.288 (4.959)	0.285 (4.905)	0.257 (4.043)
HML_t	-0.025 (1.451)	-0.023 (1.451)	0.016 (1.207)
Conditional Variables, Z_{t-1} (Dividend Yield)			
$Z_{t-1} * r_{mt}$	-	-0.048 (1.408)	-0.055 (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 III : Bootstrap Results of UK Mutual Fund Performance

Table III shows statistics for the full sample of mutual funds (including all investment objectives) for each of the three types of model selected in section IV. 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 IV. Statistical Significance of Mutual Fund Performance by Investment Objective

Table IV 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 V. Statistical Significance of Mutual Fund Performance by Fund Location

Table V 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 : $(R_i - r_f)_t = \alpha_i + \beta_{1i}(R_m - r_f)_t + \beta_{2i}SMB_t + \beta_{3i}HML_t + \epsilon_{it}$															
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 VI. 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 1a : Histograms of Bootstrap t-alpha Estimates (Upper End of the Distribution)

Figure 1a 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.

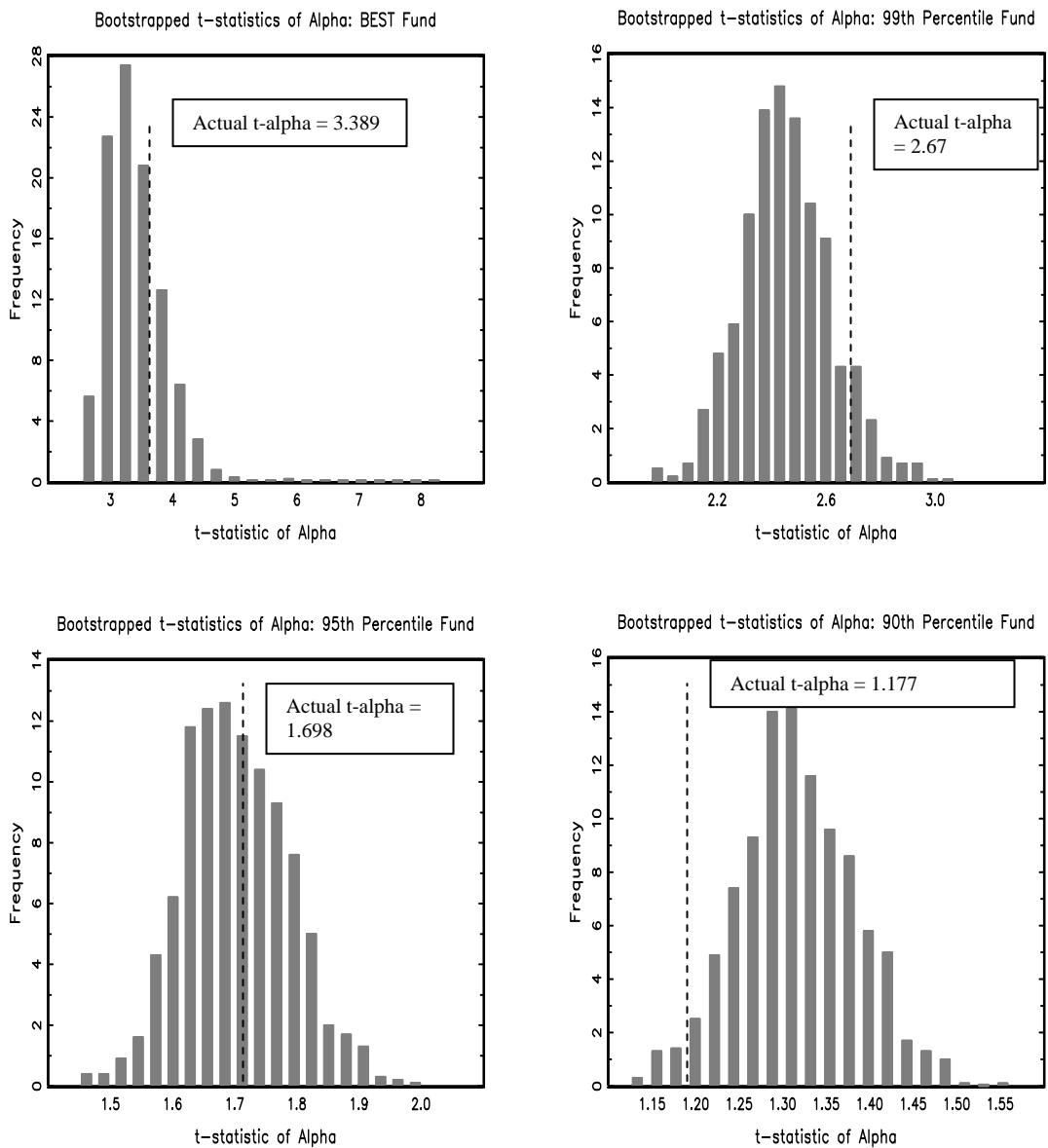


Figure 1b. Histograms of Bootstrap t-alpha Estimates (Lower End of the Distribution)

Figure 1b 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.

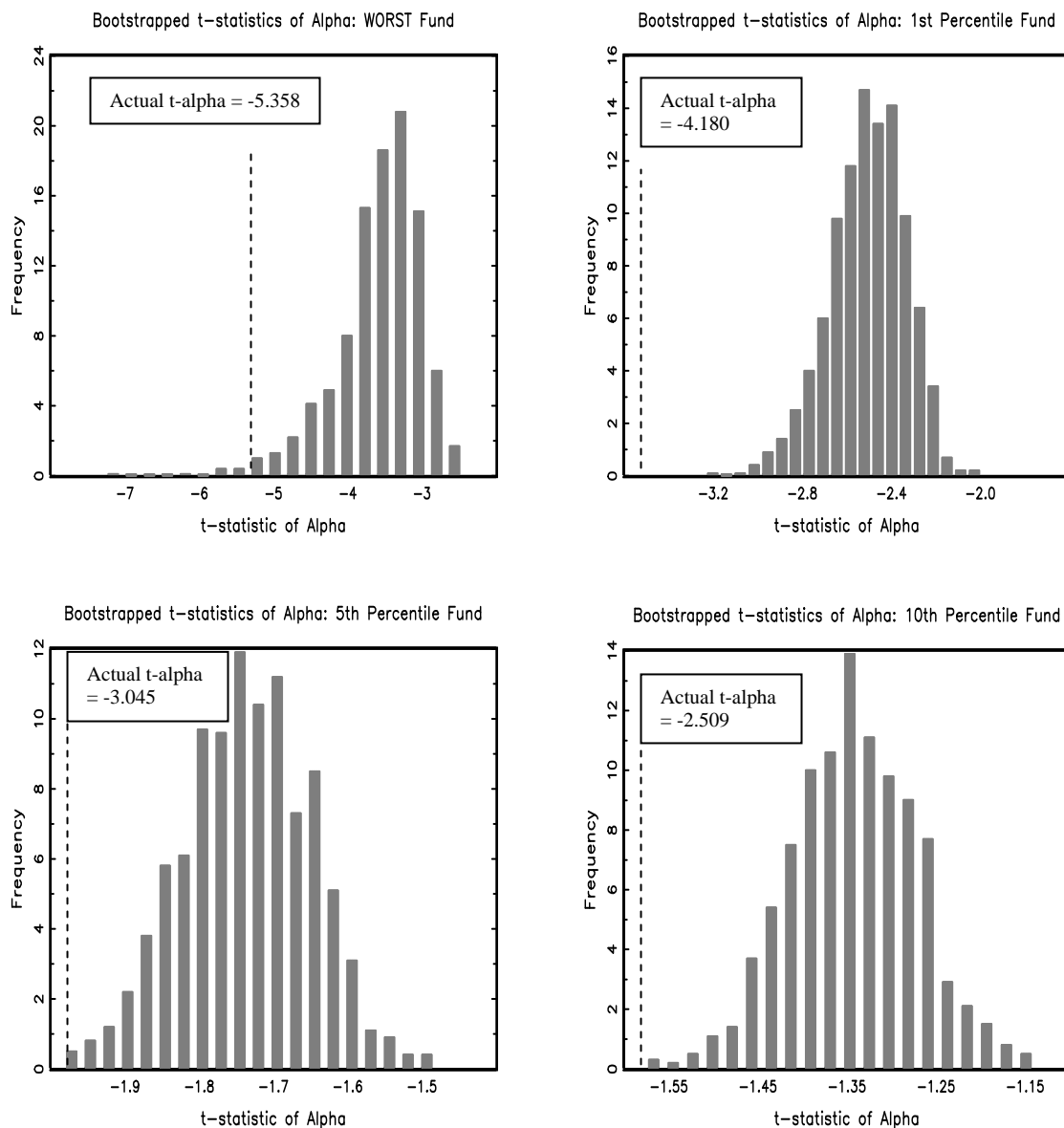


Figure 2. Kernel Density Estimates of the Actual and Bootstrap distribution of t-alphas – All Funds.

Figure 2 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.

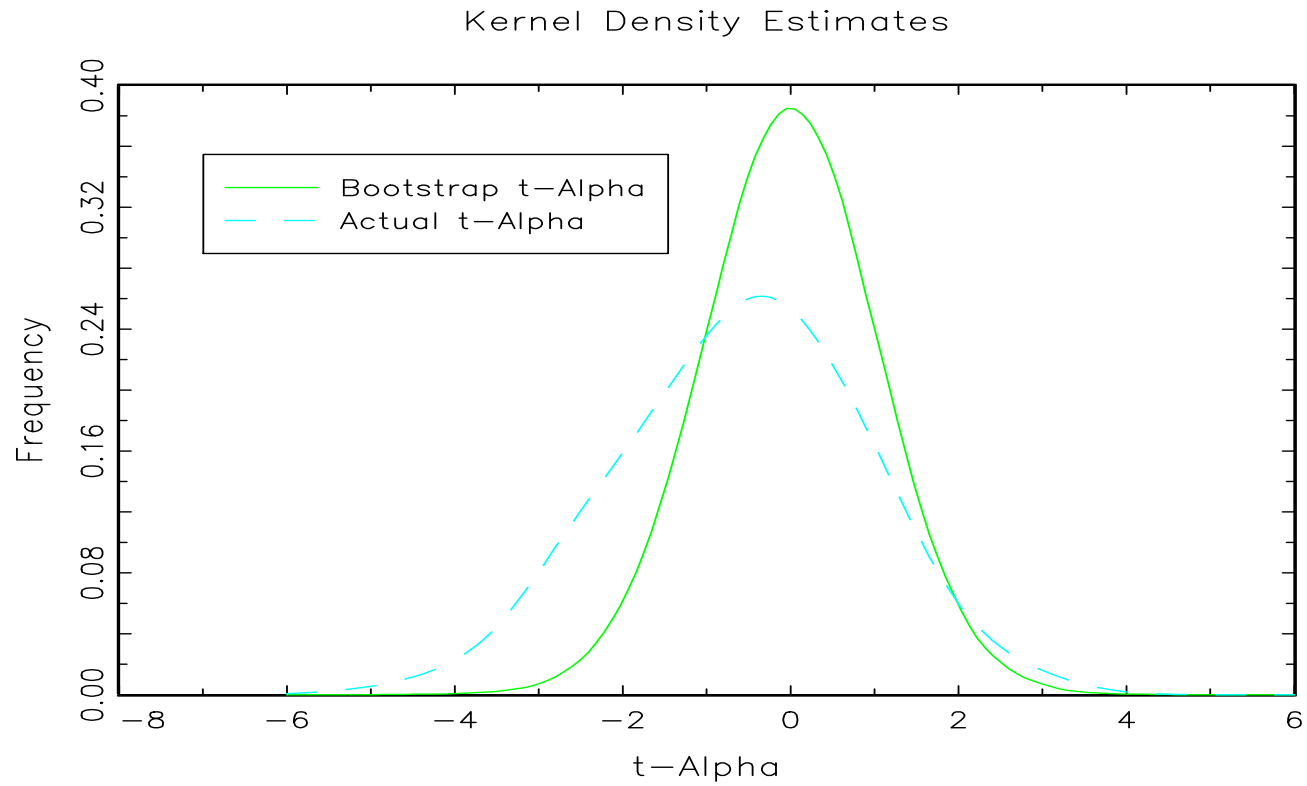
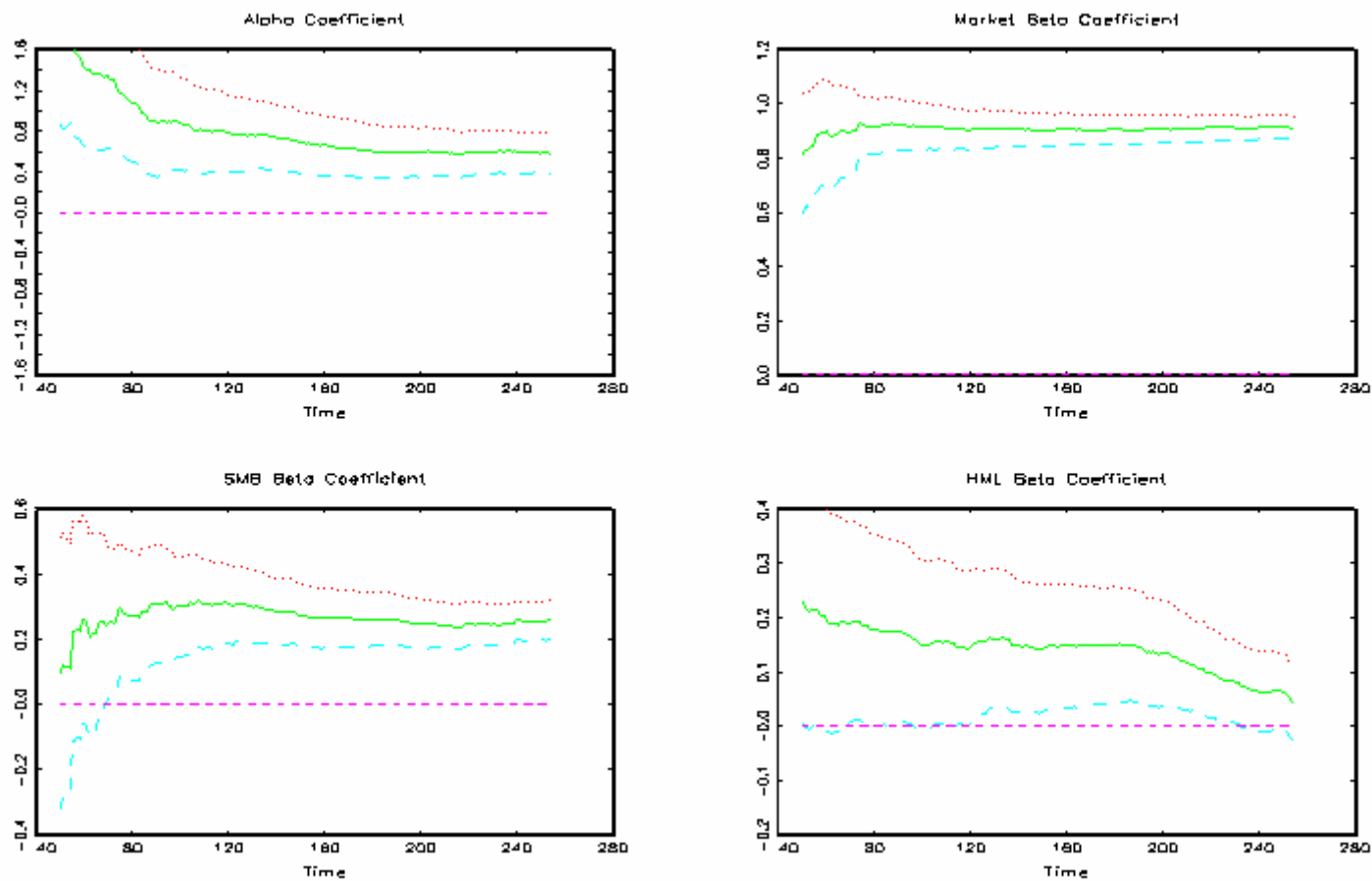
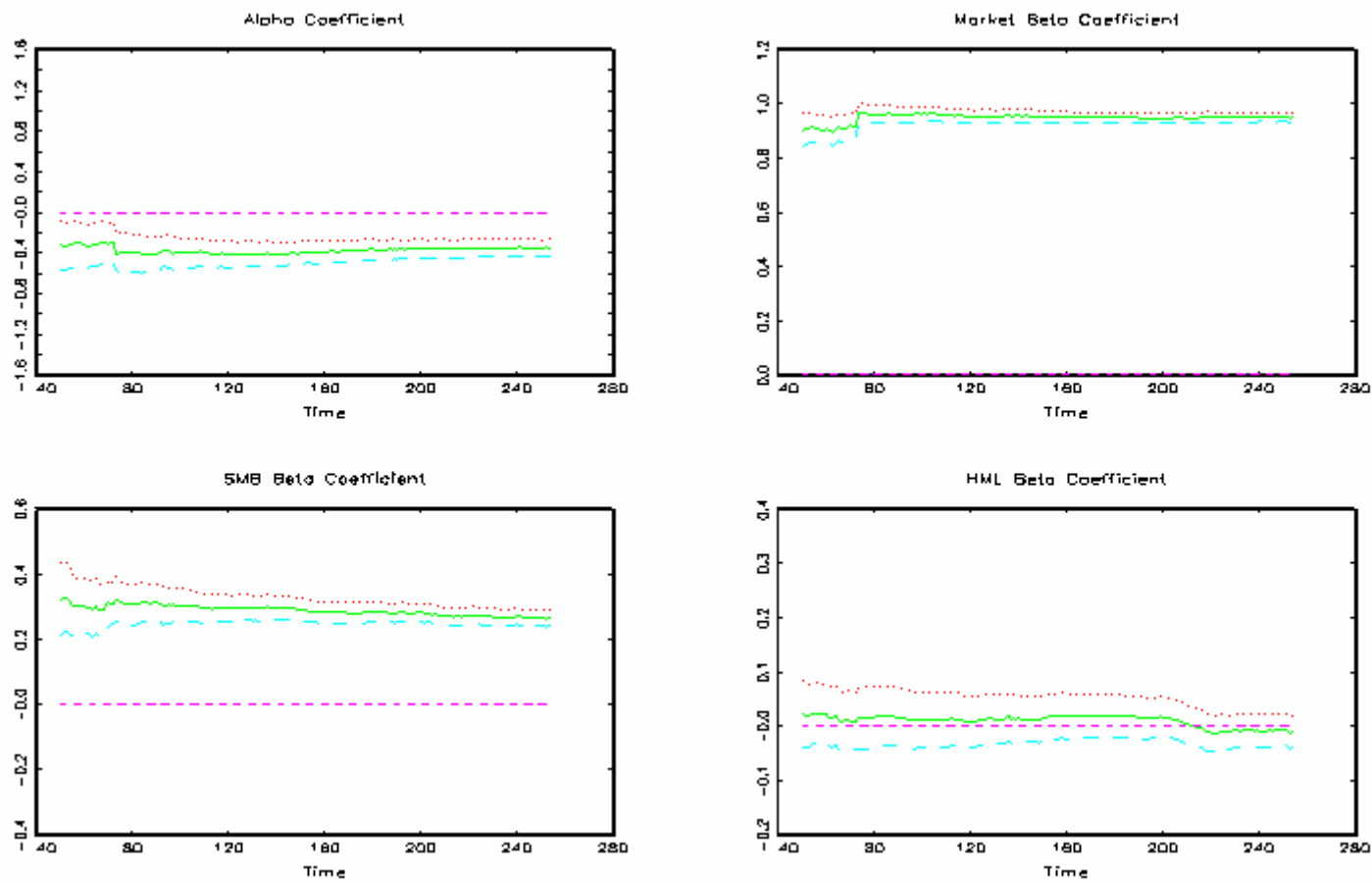


Figure 3. Recursive Estimates of Parameters (Portfolio of ‘Best 12 Funds’) : October 1981 – December 2002



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

Figure 4. 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).