

**INVESTOR HETEROGENEITY AND THE CROSS-SECTION OF
U.K. INVESTMENT TRUST PERFORMANCE**

Jonathan Fletcher and Andrew P. Marshall

University of Strathclyde

Key Words: Fund performance, investor heterogeneity, investment trusts

JEL Classification: G11, G12

Helpful comments received from an anonymous reviewer.

Address correspondence to Professor J. Fletcher, Department of Accounting and Finance, University of Strathclyde, Curran Building, 100 Cathedral Street, Glasgow, G4 0LN, United Kingdom, phone: +44 (0) 141 548 4963, fax: +44 (0) 141 552 3547, email: j.fletcher@strath.ac.uk

**INVESTOR HETEROGENEITY AND THE CROSS-SECTION OF
U.K. INVESTMENT TRUST PERFORMANCE**

ABSTRACT

We use the upper and lower bounds derived by Ferson and Lin(2010) to examine the impact of investor heterogeneity on the performance of U.K. investment trusts relative to alternative linear factor models. We find using the upper bounds that investor heterogeneity has an important impact for nearly all investment trusts. The upper bounds are large in economic terms and significantly different from zero. We find no evidence of any trusts where all investors agree on the sign of performance beyond what we expect by chance. Using the lower bound, we find that trusts with a larger disagreement about trust performance have a weaker relation between the trust premium and past Net Asset Value (NAV) performance.

I Introduction

When financial markets are incomplete, there is not a unique stochastic discount factor to evaluate fund performance¹. Ferson(2010) and Ferson and Lin(2010) point out that since marginal utility growth will vary across clients in incomplete markets, fund performance will be clientele specific. Investor heterogeneity becomes an important issue in evaluating fund performance and different clienteles of investors could value the same fund very differently². The issue of investor heterogeneity is commonly ignored in performance studies where a given benchmark model is used to evaluate the performance of managed funds. Ferson and Lin derive upper and lower bounds of the impact of investor heterogeneity on fund performance using a given benchmark model of expected returns. They find that investor heterogeneity has a substantial impact on the performance of open-end U.S. mutual funds.

We use the Ferson and Lin(2010) bounds to examine the impact of investor heterogeneity on the performance of U.K. investment trusts³ with U.K. equity objectives relative to a number of alternative benchmark models. Our study is the first one to incorporate the bounds of Ferson and Lin in the evaluation of closed-end fund performance and complements the recent study of Ferson and Lin. The performance of closed-end funds could be different from open-end funds since closed-end funds are not subject to the impact of cash flows coming into and out of the fund (Elton, Gruber and Busse(1998)). Elton et al also point out the closed-end funds might be

¹ See Chen and Knez(1996) for an overview of evaluating fund performance using the stochastic discount factor approach.

² The issue of investor heterogeneity on fund performance is also explored implicitly in the papers by Chen and Knez(1996) and Ahn, Cao and Chretien(2009).

³ Investment trusts are equivalent to closed-end U.S. mutual funds.

subject to greater agency costs since investors cannot redeem their shares at NAV. We extend the prior literature of U.K. closed-end fund performance such as Bal and Leger(1996), Bangassa(1999), and Fletcher and Marshall(2012) among others by examining the impact of investor heterogeneity on trust performance using specific benchmark models. We also examine the impact of investor heterogeneity on the relation between the trust premium and lagged Net Asset Value (NAV) performance.

There are three main findings in our study. First, we find no evidence that there are individual trusts which either add or destroy value to investors beyond what we would expect by chance. Second, we find that the upper bounds are large in economic terms and are significantly different from zero. We find that there are no trusts where all investors agree on the sign of performance beyond what we would expect by chance. Third, we find using the lower bounds that trusts with a larger disagreement about performance have a weaker relation between the trust premium and the past NAV performance. Our results suggest that investor heterogeneity has a significant impact on U.K. equity closed-end fund performance.

Our paper is organized as follows. Section II describes the research method. Section III discusses the data in our study. Section IV reports the empirical results and the final section concludes.

II Research Method

A) Fund Performance and Investor Heterogeneity

Ross(1978), Harrison and Kreps(1979), and Hansen and Richard(1987) show that if the Law of One Price (LOP) holds in financial markets, there exists a stochastic discount factor m_t such that:

$$E_{t-1}(m_t x_{it}) = p_{it-1} \quad \text{for } i = 1, \dots, N \quad (1)$$

where x_{it} is the payoff of primitive asset i at time t , p_{it-1} is the price of asset i at time $t-1$, E_{t-1} is a expected value conditional on information available to investors at time $t-1$, and N is the number of primitive assets. Equation (1) states that the conditional expected risk-adjusted payoff of asset i at time t has a price equal to p_{it-1} . Where the payoff is a gross return (1+return), the price equals 1. Where the payoff is an excess return, the price equals zero. Where financial markets satisfy the No Arbitrage (NA) restriction, m_t will be positive in every state of nature (Cochrane(2005)). In complete markets m_t will be unique.

Chen and Knez(1996) develop a framework to evaluate fund performance using the stochastic discount factor approach⁴. Define y_t as a candidate model of the stochastic discount factor. Using the excess returns on the fund, the conditional performance of the fund is given by:

$$\alpha_{pt-1} = E_{t-1}(y_t r_{pt}) \quad (2)$$

where r_{pt} is the excess return of the fund at time t . The conditional performance of the fund captures the difference between the expected risk-adjusted excess return of the fund at time t conditional on available information at time $t-1$ minus its price, which is zero.

Ferson and Lin(2010) show that the relevant stochastic discount factor to use to evaluate fund performance is the marginal utility growth of the client. When markets are incomplete, marginal utility growth can vary between clients and so performance will be clientele specific. This issue is ignored in performance studies where funds are evaluated relative to specific benchmark models and all investors are assumed to place the same value on the fund. Ferson

⁴ Aragon and Ferson(2008) provide an excellent overview of different fund performance measures and provide a survey of managed fund performance. Ferson(2010,2012) provides a review of the fund performance literature and shows how the stochastic discount factor approach can be used to unify a number of important issues in fund performance.

and Lin derive upper and lower bounds of the impact of investor heterogeneity on fund performance measures using specific benchmark models. They derive their bounds by considering a regression of the excess fund returns on a constant and the excess returns of the N primitive assets:

$$r_{pt} = a_p + \sum_{n=1}^N \beta_{pn} r_{nt} + u_{pt} \quad (3)$$

where r_{nt} is the excess return of the primitive asset n at time t , β_{pn} is the beta of fund p relative to asset n , u_{pt} is the residual term of fund p at time t with $E(u_{pt}) = 0$ and $E(u_{pt}r_{nt}) = 0$ for $n=1, \dots, N$. The intercept in regression (3) a_p is equivalent to the performance measures of Jensen(1968) and Connor and Korajczyk(1986) where the N assets are the excess returns on the factors in a given benchmark model. We can consider each benchmark model as specifying different sets of the N primitive assets. We refer to a_p relative to a given factor model as the Jensen performance of the fund.

Ferson and Lin take the unconditional expectation of equation (2) ($E(\alpha_p)$) and substitute into the regression model (3) to derive their bounds⁵. The upper bound is given by:

$$|E(\alpha_p)/E(y_t) - a_p| \leq \sigma(u_{pt})SR_{\max} \quad (4)$$

where $E(y_t)$ is the expected value of the stochastic discount factor, $\sigma(u_{pt})$ is the residual volatility from equation (3), and SR_{\max} is the maximum Sharpe ratio that can be attained by investors in the set of N primitive excess returns⁶. The lower bound is given by:

$$|E(\alpha_p)/E(y_t) - a_p| \geq |\rho_{yu}| \sigma(u_{pt})SR_{\max} \quad (5)$$

⁵ See Ferson and Lin(2010) for full details as to the derivation of their bounds.

⁶ This analysis is linked to the good deal option pricing bounds of Cochrane and Saa-Requejo(2000).

where ρ_{yu} is the correlation between the stochastic discount factor and the residuals of the fund from equation (3).

Ferson and Lin(2010) point out that when $\rho_{yu} = 0$, then the factors in the benchmark model capture the marginal utility growth of all clients and so investor heterogeneity will have no impact on fund performance. The higher the values are for $\sigma(u_{pt})$ and SR_{\max} , the greater the impact that investor heterogeneity will have fund performance relative to a specific benchmark model. Ferson and Lin point out that the use of unconditional moments is valid as long as the law of iterated expectations can be applied. They note that if investors know less information than used in a given empirical study or do not form expectations using mathematical conditional expectations, then heterogeneity in the beliefs of investors becomes an important issue. The bounds can also apply to the use of conditional moments where clients use the same public information set.

B) Estimating the Ferson and Lin(2010) Bounds

We estimate the unconditional upper bound of Ferson and Lin(2010) using the factors in a given benchmark model. The residual volatility of the fund ($\sigma(u_{pt})$) comes from the standard deviation of the residuals from equation (3). We use the adjusted maximum Sharpe ratio of Ferson and Siegel(2003) to compute SR_{\max} when $N > 1$, as in Ferson and Lin, which corrects for the upward bias in the estimated maximum Sharpe ratios using Maximum Likelihood (ML) estimation⁷. The adjusted squared Sharpe ratio is given by $[(T-N-2)/T]SR^2 - N/T$, where T is the number of observations and SR^2 is the ML estimate of the maximum squared Sharpe ratio.

We also estimate the upper bound of Ferson and Lin(2010) using conditional moments. We assume that the conditional expected excess returns of the trusts and the factor excess returns

⁷ Ferson and Siegel(2003) derive their adjusted Sharpe ratio under multivariate normality.

in a given linear factor model are a linear function of a set of lagged information variables, denoted by Z_{t-1} , and that the conditional second moments are constant. The residual volatility is then given by the standard deviation of the residuals from the regression of the trust excess returns on a constant, the lagged information variables, and the N excess factor returns (Ferson and Lin). The conditional maximum Sharpe ratio is computed each period using the fitted values from the predictive regressions of the N factor excess returns on the lagged information variables as the conditional expected excess returns and the constant conditional covariance matrix is given by the ML estimate of the residual covariance matrix from the predictive regressions. The expected value of the upper bound is then given by the residual volatility multiplied by the average conditional maximum Sharpe ratio⁸.

To evaluate statistical tests of the upper bounds, we use the bootstrap resampling approach of Fama and French(2010). Fama and French use the bootstrap resampling approach⁹ to simulate the cross-sectional distribution of fund performance under the null hypothesis of zero fund performance to compare whether the best and worst performing funds are greater than or worse than what we observe in a world of zero fund performance. We adapt the bootstrap resampling approach to generate the simulated distribution of the upper bounds of Ferson and Lin(2010) in a world of zero performance using 10,000 bootstrap samples. We estimate the standard errors of the upper bounds as the standard deviation of the simulated upper bounds across

⁸ We also correct the average conditional maximum Sharpe ratio using the adjustment in Ferson and Siegel(2003) when $N > 1$.

⁹ Full details of the bootstrap approach is included in an appendix, which is available on request. We also use the bootstrap approach to simulate the cross-sectional distribution of trust performance when there is zero performance.

the 10,000 trials. We use the standard errors of the upper bounds to compute the t -statistics of the null hypothesis that the upper bounds equal zero to examine whether investor heterogeneity is an important issue for each trust. We use a one-tail test to examine this hypothesis. If investor heterogeneity has no significant impact on the performance of a given trust, we expect the upper bound to equal zero.

As an alternative test of the impact of investor heterogeneity on trust performance, we consider the issue of whether there are any trusts for which all investors would agree on the sign of performance. Ferson and Lin(2010) point out that all investors will only agree on the sign of performance when the absolute value of a_p is larger than the upper bound. We compute the cross-sectional distribution of $|a_p| - \text{Upper Bound}$. Investors will only agree on the sign of performance when $|a_p| - \text{Upper Bound} > 0$. We use the bootstrap resampling approach of Fama and French(2010) and estimate the cross-sectional simulated distribution of $|a_p| - \text{Upper Bound}$ in a world of zero performance. We compare the actual cross-sectional distribution with the simulated distribution to examine the null hypothesis that $|a_p| - \text{Upper Bound} = 0$ at selected percentiles by calculating the proportion of times out of the 10,000 trials that the simulated $|a_p| - \text{Upper Bound}$ lie below the actual $|a_p| - \text{Upper Bound}$. If all investors agree on the sign of performance, for the percentiles where $|a_p| - \text{Upper Bound} > 0$, we expect the proportion of simulated $|a_p| - \text{Upper Bound}$ values lower than the actual value to be close to 1.

We use the lower bound of Ferson and Lin(2010) to examine whether investor heterogeneity has any impact on the relation between the premium of the trust and past NAV performance given the performance theory of the discount (e.g. Berk and Stanton(2007))¹⁰. Ferson and Lin use the lower bound to examine whether investor heterogeneity has any impact

¹⁰ We are grateful to the reviewer suggesting that we explore this issue.

on the relation between fund flows and past performance for open-end mutual funds. The performance theory of the discount predicts that there is a positive relation between the trust premium and lagged NAV performance. We conjecture that this relation would be weaker for those trusts where there is a greater disagreement among investors over trust performance.

We use a measure of disagreement (Disagree_p) motivated by the first disagreement measure of Ferson and Lin(2010). Ferson and Lin estimate their measure for each fund as $\sigma(u_{pt}) * \sigma(|\rho(u_{pt}, G)|)$ where $\rho(u_{pt}, G)$ is the correlation between the residuals of the fund and the electricity consumption growth of a given U.S. state. The state electricity consumption growth is used to proxy for marginal utility growth of different investors (see Da and Yun(2010)). The $\sigma(|\rho(u_{pt}, G)|)$ terms captures the variation across investors of the absolute correlations between the fund's residuals and marginal utility growth. The larger the residual volatility of the fund and the greater the variation in absolute correlations across investors, the larger the Disagree_p measure.

Given the need to estimate the marginal utility growth of different investors and the unavailability of disaggregated electricity consumption growth in the U.K. for our whole sample period, we adopt a different approach in our study. We back out proxies of marginal utility growth using the Hansen and Jagannathan(1991) NA volatility bounds. The NA volatility bounds are the set of minimum variance nonnegative (admissible) stochastic discount factors that price a set of test assets. Chen and Knez(1996) point out that a NA admissible stochastic discount factor can be viewed as the marginal utility growth of an investor¹¹. By selecting

¹¹ Ahn et al(2009) extend the work of Chen and Knez(1996) by deriving the performance bounds of a fund implied by a set of NA admissible stochastic discount factors. They interpret the upper (lower) performance bound as the performance value of the fund from the perspective of the

different NA admissible stochastic discount factors from the NA volatility bounds, we are implicitly capturing the marginal utility growth of different investors.

We estimate the NA volatility bounds using the gross returns of a set of payoffs in U.K. stock returns across the whole sample period. We select 25 NA admissible stochastic discount factors that lie on the bounds by setting $E(m)$ initially to 0.975 and increasing $E(m)$ by 0.001 at a time. For each trust, we calculate the correlations between u_{pt} and each of the NA admissible stochastic discount factors and estimate the corresponding $Disagree_p$ measure. Our approach is admittedly ad hoc and could be sensitive to the set of payoffs we use to calculate the bounds and number of admissible stochastic discount factors selected.

We use the Fama and MacBeth(1973) cross-sectional regression approach to examine the impact of investor heterogeneity on the relation between the trust premium and past NAV performance. Each year between 1993 and 2010, we run the cross-sectional regression:

$$Prem_{pt} = \gamma_{0t} + \gamma_{1t}a_{pt} + \gamma_{2t}a_{pt} * Disagree_p + Control\ Variables_{pt} + e_{pt} \quad (6)$$

where $Prem_{pt}$ is the average monthly premium during year t for trust p , and a_{pt} is the past NAV performance relative to a linear factor model for trust p at the start of year t . We measure the premium as the ratio of the share price to NAV at the end of each month. We estimate the past performance for all trusts with continuous NAV excess returns during the past 36 months. We assume that the $Disagree_p$ measure is fixed for each trust and we use both the actual values of $Disagree_p$ measure and construct dummy variables of the $Disagree_p$ measure as in Ferson and Lin(2010). We construct two dummy variables, $Disagree_{pH}$ which equals 1 if the trust is in the largest 1/3 of $Disagree_p$ measures across trusts and 0 otherwise, and $Disagree_{pM}$ which equals 1 if

investor clientele most (least) favorable to the fund. Fletcher and Marshall(2012) use the NA performance bounds of Ahn et al(2009) to evaluate the performance of U.K. investment trusts.

the trust is in the middle 1/3 of Disagree_p measures across trusts and 0 otherwise. When we use dummy variables for the Disagree_p measure, there will be two interaction independent variables in equation (6). The time-series average of the annual γ coefficients are used as the final estimates and the t -statistics are calculated from the time-series of the annual γ_t coefficients. The use of the Fama and MacBeth approach controls for the impact of residual cross-correlations across trusts. If investor heterogeneity has an impact on the relation between the premium and past NAV performance, we expect the average values of the slope coefficients on the interaction variables (i.e. γ_2) to be significantly negative.

We use a selection of control variables in equation (6) using a subset of similar variables in Khorana, Servaes and Wedge(2009). We use the log of the total NAV of the trust, annual DY, and age of the trust at the start of year t . We calculate the age of the trust as in Ferson and Lin(2010) as the log of the number of months since the trust began. We do not include a control variable for the expense ratio as Khorana et al find no relation between fund expenses and the premium (see also Malkiel(1977)).

IV Data

All of the data for this study is collected from the London Share Price Database (LSPD) from the London Business School unless otherwise specified. Full details of our sample of trusts, the construction of the factors in the linear factor models, and lagged information variables used in our study are included in an appendix, which is available on request.

A) Sample of Investment Trusts

Our sample of funds includes all U.K. investment trusts with a U.K. equity objective between January 1990 and December 2010. The U.K. equity sectors include the U.K. Growth, U.K. Growth and Income, U.K. Small Companies, and U.K. High Income investment sectors. The investment sector information for each trust is collected at the start of each year from Money

Management¹² between 1990 and 2009 and the Association of Investment Companies (AIC) web site for the last year. There are 221 investment trusts in our sample. Our sample of trusts should be relatively free of survivorship bias (Brown, Goetzmann, Ibbotson and Ross(1992)). However our results using individual trusts can be affected by the reverse survivorship bias of Linnainmaa(2012). Linnainmaa argues that reverse survivorship bias leads to an understatement of true performance when funds disappear because of poor performance where the poor performance is in part due to negative idiosyncratic shocks.

For each trust, we collect the stock returns of the trusts and calculate excess returns using the monthly return of the one-month U.K. Treasury Bill (collected from LSPD and Datastream) as the risk-free asset. Aragon and Ferson(2008) interpret the use of fund returns which are gross of all trading costs and expenses as capturing the performance ability of funds and where fund returns are net of all costs and expenses as capturing the value added to investors. The use of stock returns for investment trusts does capture the value added of the trust to investors. The value added depends not only on performance ability, costs, expenses, but also on the behavior of the trust premium¹³. We also calculate the NAV returns of the trusts using the monthly NAV

¹² The investment sectors have changed names over the years. The four sectors are the current names of the U.K. investment sectors as at the end of the sample period. In the early part of the sample period, there was a U.K. General sector. We allocate trusts in the U.K. General sector to the U.K. Growth sector since most trusts transferred to this sector when the classifications changed.

¹³ See Dimson and Minio-Paluello(2002) for a review of the alternative explanations of the closed-end fund discount. Recent studies by Berk and Stanton(2007) and Cherkes, Sagi and Stanton(2009) develop theories of the fund premium in relation to expectations about future

of the trusts collected from Datastream and the dividends paid by the trusts. We estimate the performance and the lower and upper bounds of Ferson and Lin(2010) for all trusts with greater than 24 return observations, which creates a look-ahead bias (Carhart, Carpenter, Lynch and Musto(2002)).

B) Linear Factor Models

We evaluate the performance of the investment trusts using five linear factor models. We use domestic factor models since the trusts are U.K. equity funds and Griffin(2002) finds that a local version of the Fama and French(1993) model performs better than the global version of the model. The models include:

1. CAPM

This model is a single-factor model that uses the excess returns (over the one-month Treasury Bill return) of the U.K. stock market index (Market) as the proxy for aggregate wealth.

2. Fama and French(1993) (FF)

The FF model is a three-factor model. The factors are the excess return on the market index and two zero-cost portfolios that capture the size (SMB) and value/growth (HML) effects in stock returns. We use the price-to-book (PB) ratio from Datastream to capture the value/growth effects.

3. Carhart(1997)

The Carhart model is a four-factor model. The factors are the three factors in the FF model and a zero-cost portfolio that captures the momentum effect (WML) in stock returns.

managerial performance ability or to the liquidity benefits provided by the funds. Ramadorai(2012) provides support for these theories in explaining the closed hedge fund premium.

4. Four-index model (4-index)

This model is a four-factor model and is motivated by the four-index model in Cremers, Petajisto and Zitzewitz(2010). Cremers et al advocate the use of index-based models to capture the size and value/growth effects in stock returns. The factors include the excess returns (over the one-month Treasury Bill return) on the largest 100 stocks (Large), the difference in returns between small stocks and large stocks (S-L), the difference in returns between low PB stocks and high PB stocks across all companies (AHML), and WML.

5. Seven-index model (7-index)

This model is a seven-factor model and is motivated by the seven-index model in Cremers et al(2010). The factors include the excess returns on the largest 100 stocks, the difference in returns between small stocks and mid-cap stocks (S-M), the difference in returns between mid-cap stocks and large stocks (M-L), the difference in returns between low PB stocks and high PB stocks across large companies (LHML), the difference in returns between low PB stocks and high PB stocks across mid-cap companies (MHML), the difference in returns between low PB stocks and high PB stocks across small companies (SHML), and WML.

Table 1 present summary statistics of the monthly excess returns of the sample of investment trusts and the factors in the linear factor models. The summary statistics include the mean, standard deviation, minimum, and maximum of monthly excess returns (%). In panel A of the table, all trusts with greater than 24 stock return observations are sorted separately by the statistic in the column and various percentiles of the cross-sectional distribution across individual trusts are reported. Panel B reports the summary statistics for the factors in the linear factor models.

Table 1 here

Panel A of Table 1 shows that there is a wide spread in the cross-section of the summary statistics of the stock excess returns of the individual trusts. The average excess returns range between -0.774% for the bottom 10% of trusts and 1.050% for the top 10% of trusts. The volatility of excess returns ranges between 5.121% for the bottom 10% of trusts and 18.491% for the top 10% of trusts. The median average excess return is larger than the average excess market return in panel B but the median volatility is considerably larger than the volatility of the market excess returns. The summary statistics of the individual trusts suggest that there is substantial heterogeneity across trusts.

The WML factor has the largest average excess returns of the factors in panel B of Table 1. There is a large significant positive average excess return on the WML factor of 1.015% which highlights the strong momentum effect in U.K. stock returns during this sample period. The only other factors with significant positive average excess returns are the HML factor and the HML (Small) factor, which suggests that the value/growth effect is strongest in smaller companies during our sample period. There is no significant size effect as the average excess returns on the SMB, S-L, S-M, and M-L factors are tiny. The average excess market returns are also not statistically significant due to impact of the large market downturn in 2007 and 2008.

C) Test Assets Used To Form NA Admissible Stochastic Discount Factors and Lagged Information Variables

We use the gross returns of a set of nine payoffs when forming the minimum variance NA admissible stochastic discount factors that we use to proxy for the marginal utility growth of different clients to estimate the lower bound of Ferson and Lin(2010). We use the six size/PB

index portfolios used in the formation of the 7-index model of Cremers et al(2010), the Winners and Losers portfolios used to form the WML factor, and the Datastream U.K. government bond index (All Lives). To estimate the upper bound of Ferson and Lin(2010) using a model of conditional moments, we use a similar set of lagged information variables as in Ferson and Lin(2010). We use the lagged annualized dividend yield on the market index, the lagged return on the three-month U.K. Treasury Bill, the lagged term spread, and the lagged default spread¹⁴.

IV Empirical Results

We begin our empirical analysis by examining the cross-sectional distribution in individual trust performance. Table 2 reports percentiles from the cross-sectional distribution of individual trust performance for all trusts. The table reports the percentiles for the Jensen performance (a_p), the standard error of performance ($SE(a_p)$), and the volatility of the residuals ($\sigma(u_{pt})$) where the trusts are sorted separately by the statistic in the column. The final column of Table 2 reports the performance and t -statistic in parentheses of the value weighted portfolio of all trusts ($a_p(VW)$) relative to each model. The standard errors and t -statistics in Table 2 are corrected for the effects of heteroskedasticity and serial correlation using the automatic lag selection (without prewhitening) of Newey and West(1994).

Table 2 here

¹⁴ The motivation for including these information variables stems from the large literature on predictable stock returns. See Lettau and Ludvigson(2010) for an excellent review of this literature.

Table 2 shows that there is a large cross-sectional spread in individual trust performance. This result holds for all linear factor models. The median a_p is positive for all models except for the FF model where the median negative performance is tiny. The average trust performance is positive for the CAPM, 4-index, and 7-index models and negative for the FF and Carhart models. The average performance is affected by the extreme outliers of performance by some trusts. The performance of the portfolio of all trusts is small and close to zero and none are statistically significant for all factor models. This result suggests that the performance of the trusts, in aggregate, is neutral and neither adds or destroys value to investors.

To explore the impact of the benchmark model on performance in more detail, for each trust we calculate the range between the largest and smallest a_p across the five linear factor models as in Ferson and Lin(2010). We then compute the percentiles of the cross-sectional distribution of the range in performance across the factor models. The median range in performance across models is 0.541% and the interquartile range of the variation in performance is between 0.323% and 1.254%. The range in performance suggests that the benchmark model has a significant impact on individual trust performance and is similar to prior studies such as Lehmann and Modest(1987), Grinblatt and Titman(1994), Fletcher and Forbes(2002), and Ferson and Lin(2010) on open-end funds.

The standard errors of the estimated performance of the individual trusts in Table 2 are large in many cases and show a great deal of variation across trusts. The median standard error of a_p varies between 0.488% (FF) and 0.554% (CAPM) using excess stock returns. The large standard errors suggest that the performance of the trusts will be estimated imprecisely and will be difficult to detect either significant superior or inferior performance. The magnitude of the median standard error is similar in size to the median range in performance across models, which

is similar result to Ferson and Lin(2010) although the estimated medians are larger for closed-end funds here compared to the open-end mutual funds of Ferson and Lin.

Table 2 also shows that there is a wide spread in residual volatility across the individual trusts for each factor model. The median residual volatility ranges between 4.978% (7-index) and 5.981 (CAPM). The spread between the bottom 10% of trusts and the top 10% of trusts in residual volatility is 3.375% and 17.345% for the CAPM and 3.051% and 16.635% for the 7-index model. This result suggests that a large proportion of the volatility of the excess stock returns of the trusts are not captured by the benchmark models. The large spread in residual volatility across trusts suggests that the impact of investor heterogeneity on fund performance will vary across trusts (Ferson and Lin(2010)).

We next use the bootstrap resampling approach of Fama and French(2010) to examine whether the best and worst performing investment trusts is due to skill or luck. Table 3 reports the results for the linear factor models. The table reports selected percentiles from the cross-sectional distribution of actual $t(a_p)$ statistics¹⁵. In parentheses below are the average simulated (across 10,000 trials) $t(a_p)$ statistic in the () brackets under the null hypothesis of zero trust performance and the proportion (%) of simulated $t(a_p)$ statistics that lie below the corresponding actual $t(a_p)$ statistic across the 10,000 trials in the [] brackets.

Table 3 here

¹⁵ We use the $t(a_p)$ statistic, rather than a_p , since it corrects for differences in standard errors across trusts due to different residual volatilities, and a different number of return observations (see Fama and French(2010)).

Table 3 shows that there is no evidence of significant superior performance by trusts relative to any factor model. For the top 1% of trusts, the actual $t(a_p)$ statistic is below the average simulated $t(a_p)$ statistic. For the top 10% of trusts, the actual $t(a_p)$ statistic only lies above the average simulated $t(a_p)$ statistic for the 4-index model. In the CAPM, FF, and Carhart models, the simulated $t(a_p)$ statistic lies above the actual $t(a_p)$ statistic in the vast majority of trials for both the top 10% and 1% of trusts. In the 4-index and 7-index models, the simulated $t(a_p)$ statistic lies above the actual $t(a_p)$ statistic in the vast majority of trials for the top 1% of trusts. For the top 10% of trusts relative to the 4-index and 7-index models, the actual $t(a_p)$ statistic lies above the simulated $t(a_p)$ statistic in the majority of trials but is well below levels which are deemed to produce significant superior performance. These results suggest that the best performing investment trusts do not add significant value to investors beyond what we would expect by chance. This result is similar to Fama and French(2010) who find that the best performing U.S. open-end mutual funds do not add significant value using net fund returns.

On a more positive note, we find no evidence of significant inferior performance by the worst performing trusts beyond what we would expect by chance. For the bottom 10% and bottom 1% of trusts, the average simulated $t(a_p)$ statistic lies below the actual $t(a_p)$ statistic. For the CAPM, 4-index, and 7-index models, the simulated $t(a_p)$ statistic lies below the actual $t(a_p)$ statistic in the vast majority of trials at the bottom 25%, 10%, and 1% of trusts. For the FF and Carhart models, the simulated $t(a_p)$ statistic lies below the actual $t(a_p)$ statistic in the majority of trials at the bottom 25%, 10%, and 1% of trusts with the exception of the FF model at the bottom 10% of trusts. The results suggest that the worst performing trusts do not destroy value for investors beyond what we would expect in a world of zero performance. This finding stands in

contrast to that observed for open-end U.S. mutual funds in Fama and French(2010). Fama and French find that the worst performing funds have significant inferior performance.

The results in Table 3 suggests that individual trusts neither add or destroy value to investors when we ignore the issue of investor heterogeneity on the performance of the individual trusts¹⁶. However when we recognise the issue of investor heterogeneity it is possible that different investors will value the trusts differently from what we observe in Tables 2 and 3. We examine the impact of investor heterogeneity on the performance of the trusts using the different benchmark models. We estimate the upper bounds of Ferson and Lin(2010) in equation (4) using unconditional moments. Panel A of Table 4 reports percentiles from the cross-sectional distribution of upper bounds of investor heterogeneity on fund performance. The SR row at the bottom of panel A is the maximum (adjusted) Sharpe ratio available from the set of the N primitive excess returns in a given benchmark model. In panel B of the table, we report percentiles from the cross-sectional distribution of t -statistics of the null hypothesis that the upper bounds equal zero. N is the number of trusts where the t -statistic is significant at the 5% level using a one-tail test. Panel C of the table reports the percentiles from the cross-sectional distribution of the actual values of $|a_p| - \text{Upper Bound}$. In parentheses below is the proportion (%) of times that the simulated values of $|a_p| - \text{Upper Bound}$ lie below the actual values of $|a_p| - \text{Upper Bound}$.

Table 4 here

¹⁶ The lack of funds with significant positive alphas might be due to the reverse survivorship bias of Linnainmaa(2012).

Panel A of Table 4 shows that investor heterogeneity has a large impact on the performance measures using linear factor models. Different investors are likely to evaluate the trusts very differently from what we observe in Tables 2 and 3. The median upper bound in panel A of Table 5 ranges between 0.420% (CAPM) and 2.048% (Carhart). The median upper bounds are a lot larger for the Carhart, 4-index, and 7-index models compared to the CAPM and FF models. This result stems from the fact that the maximum Sharpe ratio is a lot larger for these three models due to the inclusion of the WML factor. For the bottom 1% of trusts relative to the multifactor models, the upper bound is 0.345% (FF), 0.807% (Carhart), 0.617% (4-index), and 0.694% (7-index). Even for the FF model, the upper bound for the bottom 1% of trusts is 0.345%. These results suggest that investor heterogeneity is likely to be an important issue for nearly every investment trust using the multifactor models.

We can compare the point estimates of the upper bounds in panel A of Table 4 to the uncertainty of the estimated performance measures using the linear factor models ($SE(a_p)$) and the impact of the benchmark model in fund performance as in Ferson and Lin(2010). The median standard error of a_p is higher than the median upper bound for the CAPM model but the median upper bound is considerably larger than the median $SE(a_p)$ for the multifactor models. The median range (best and worst performance) in a_p across the 5 benchmark models is 0.541%, which is larger than median upper bound of the CAPM model but considerably lower than the median upper bounds for the multifactor models. These results suggest that investor heterogeneity is more important than either the uncertainty in the estimated a_p and the choice of the benchmark model when using the upper bounds from the multifactor models. The impact of investor heterogeneity on investment trust performance is greater than observed in open-end U.S. mutual funds in Ferson and Lin(2010). The difference of our findings with Ferson and Lin stems

from the higher residual volatility of our sample of investment trusts compared to that observed in Ferson and Lin.

The point estimates of the upper bounds in panel A of Table 4 suggest that investor heterogeneity will be an important issue for nearly all trusts. This finding is confirmed in the statistical tests in panels B and C of Table 4. In panel B, for the multifactor models nearly all trusts have upper bounds that are significantly different from zero. The proportion of trusts with a significant t -statistic rejecting the null hypothesis of a zero upper bound is greater than or equal to 98.4% for the multifactor models. Even for the bottom 10% of trusts, the t -statistics are large especially for the Carhart, 4-index, and 7-index models. For the CAPM model, we are not able to reject the null hypothesis that the upper bound equals zero for any trust. This result stems from the smaller maximum Sharpe ratio compared to the other benchmark models. The lack of statistical significance of the upper bounds for the CAPM would question the importance of investor heterogeneity for this model.

In panel C of Table 4, we provide a different perspective on investor heterogeneity and consider whether there are any trusts where investors will agree on the sign of the trust performance. It is only for the CAPM, where for the top 25% of trusts the absolute performance of the trust exceeds that of the upper bound and investors would agree on the sign of the performance. For the FF model, it is only for top 10% of trusts where the absolute performance of the trust exceeds the upper bound. For the Carhart, 4-index, and 7-index models, it is only for the top 1% of trusts where the absolute performance exceeds that of the upper bound. These results suggest for the multifactor models, that it is only for a small minority of trusts where investors will actually agree on the sign of the performance when looking at the actual values of $|a_p| - \text{Upper Bound}$. However when we consider the simulation results in parentheses, none of

the cases where the absolute performance exceeds the upper bound are greater than what you expect by chance. The only case near to levels of statistical significance is for the FF model where for the top 1% of trusts the actual $|a_p| - \text{Upper Bound}$ exceeds the simulated $|a_p| - \text{Upper Bound}$ in 87.9% of trials.

Table 4 suggests that with the exception of the CAPM, the upper bounds are significantly different from zero for nearly all trusts. There is no evidence that investors will agree on the sign of performance for any trust beyond what we would expect by chance. These results suggest that investor heterogeneity will be an important issue for investment trusts¹⁷. We next examine the impact of using conditional moments to compute the upper bounds. Ferson and Lin(2010) point out that the use of conditional moments can have two effects on the upper bounds. The first effect is that the funds can have lower residual volatility, which would lower the upper bound. The second effect is a higher expected conditional Sharpe ratio using conditional moments, which would lead to a higher upper bound. The impact of conditional moments on the upper bound will depend upon which effect dominates. Table 5 reports the same results as Table 4. The SR row at the bottom of panel A is the maximum (adjusted) average conditional Sharpe ratio available from the set of the N primitive excess returns in a given benchmark model.

Table 5 here

¹⁷ In unreported tests (available on request), we estimate the upper bounds using the NAV excess returns of the trusts. We find similar results to Table 4 except the magnitude of the bounds is lower.

The bottom row of panel A of Table 5 shows that there is a sharp rise in the average conditional Sharpe ratio in the N primitive excess returns for each benchmark model when we incorporate conditioning information compared to the unconditional Sharpe ratios in Table 4. Allowing for conditioning information leads to a more favorable risk/return trade-off and is similar to Ferson and Siegel(2009) among others. This increase shows that the impact on the Sharpe ratio of using conditioning information is more important than the impact on residual volatility, which leads to the expected upper bounds in panel A of Table 5 being a lot larger than that in Table 4.

The median expected upper bounds in panel A of Table 5 range between 1.295% (CAPM) and 3.111% (7-index). The impact of investor heterogeneity is again largest for the Carhart, 4-index, and 7-index models. Investor heterogeneity will be an important issue for all investment trusts as even the bottom 1% of trusts has an expected upper bound of 0.466% (CAPM), 0.907% (FF), 1.239% (Carhart), 1.032% (4-index), and 1.220% (7-index).

The point estimates in panel A of Table 5 suggest that investor heterogeneity has a greater impact when using the model of conditional moments compared to panel A of Table 4. This finding also holds in the statistical tests of panels B and C of Table 5. In panel B of Table 5, nearly all of the trusts have significant upper bounds across all five linear factor models, even for the CAPM. We can reject the null hypothesis that the upper bounds of the trusts equal zero for greater than or equal to 98.9% of trusts. The magnitude of the t -statistics are larger in panel B compared to the corresponding t -statistics in Table 4.

In panel C of Table 5, it is only for the top 1% of trusts where investors agree on the sign of performance for four of the five factor models using the point estimates of $|a_p|$ - Upper Bound. For the Carhart model, even for the top 1% of trusts investors do not agree on the sign of

performance. The simulation results in parentheses suggest that for the top 1% of trusts relative to the CAPM, 4-index, and 7-index models, the actual value of $|a_p|$ - Upper Bound is not greater than we would expect by chance. It is only for the FF model, where the actual value of $|a_p|$ - Upper Bound for the top 1% of trusts is close to levels of statistical significance where the actual $|a_p|$ - Upper Bound exceeds the simulated $|a_p|$ - Upper Bound in 91.94% of trials. The findings in panel C of Tables 4 and 5 link in with the results of Table 3. Given that in Table 3, we find no evidence of trusts with significant superior performance and in panel C of Tables 4 and 5 we find no trust which generates performance above the upper bound beyond what we expect by chance, then there is no trust about which all investors can agree that the trust adds value.

The results in Table 5 suggests that using a model of conditional moments to estimate the upper bounds implies that investor heterogeneity will have an even greater impact on evaluating investment trust performance. This result is similar to Ferson and Lin(2010) for open-end funds. However the magnitude of our bounds are greater due to the higher residual volatility of our sample of closed-end funds. The importance of investor heterogeneity on performance is also consistent with Ahn et al(2009).

The final issue we address is to examine whether investor heterogeneity has any impact on the relation between the trust premium on the past NAV performance. Table 6 reports the results from the Fama and MacBeth(1973) cross-sectional regressions of equation (6). The table reports the average coefficients from the annual cross-sectional regressions and the t -statistics in parentheses. The t -statistics are corrected for the effects of heteroskedasticity and serial correlation using the automatic lag selection (without prewhitening) method of Newey and West(1994). The final row of each panel is the average adjusted R^2 from the annual cross-sectional regressions. In panel A of the table, we report the cross-sectional regression results

excluding the Disagree_p interaction variables. Panel B (C) reports the results using the actual values of the Disagree_p measure (Disagree_{pH} and Disagree_{pM} dummy variables).

Table 6 here

Panel A of Table 6 shows that there is a positive relation between the average trust premium and the lag NAV performance. The positive relation is statistically significant for the Carhart and 7-index models. This positive relation provides some support for the performance theory of the discount and is similar to Khorana et al(2009) for U.S. closed-end funds and Dimson and Minio-Paluello(2001) for U.K. funds, although both studies use different measures of past NAV performance. Two out of the three control variables have a significant relation with the average premium of the trust. There is a significant positive relation between the average premium and the size of the trust (using total NAV) and between the average premium and the DY of the trust. Larger trusts and trusts with higher DY have larger average premiums. The relation between size and the trust premium differs from Khorana et al who find a significant negative relation.

Panel B of Table 6 shows that when we include the Disagree_p interaction variable in the cross-sectional regressions, there is now a significant positive relation between the average premium and lag NAV performance across all factor models. There is a negative coefficient on the Disagree_p interaction variable across all models, which suggests that trusts with a greater disagreement about performance have a weaker relation between the average premium and lag NAV performance. The negative coefficient is statistically significant for the CAPM and 4-

index models at the 10% significance level. The size and DY control variables remain significant in panel B of Table 6.

We can use the magnitude of the slope coefficients on a_p and $a_p * \text{Disagree}_p$ to examine the impact of investor heterogeneity on the relation between the average trust premium and past NAV performance other things being equal. For a 1% change in the lag NAV performance, the change in the average monthly premium will range (for the trusts with the minimum and maximum levels of disagreement) between 0.195 and -1.027 (CAPM), 0.144 and -0.618 (FF), 0.107 and -0.112 (Carhart), 0.115 and -0.671 (4-index), and 0.156 and -0.448 (7-index). The range in the change in the average monthly premium in response to the lag NAV performance is extremely wide and suggests that the magnitude of the slope coefficient on the Disagree_p interaction variable is economically significant.

Using the dummy variables to capture the disagreement about trust performance in panel C of Table 6 confirms that investor heterogeneity has an impact on the relation between the average premium and the lag NAV performance. There is a significant positive relation between the average premium and the lag NAV performance for the Carhart, and 7-index models. The magnitude of the coefficient on a_p captures the relation between the average premium and the lag NAV performance, other things being equal, for trusts with the smallest levels of disagreement. For these trusts, a 1% change in the lag NAV performance leads to a change in the average premium ranging between 0.071 (CAPM) and 0.140 (Carhart).

There is a negative coefficient on the $a_p * \text{Disagree}_{pH}$ variable. The change in the average premium to a 1% change in the lag NAV performance, other things being equal, for the trusts with the largest disagreement range between -0.122 (4-index) and 0.074 (7-index). Trusts with the largest disagreement either have a negative change in the premium or a smaller positive

change in the premium in response to the lag NAV performance. The magnitude of the slope coefficients are large in economic terms for all models except the 7-index model.

The coefficients on the $a_p * \text{Disagree}_{pM}$ variable are either negative or positive depending upon the benchmark model used. For trusts in the Medium disagreement group, the change in the average premium for a 1% change in the lag NAV performance ranges between 0.018 (FF) and 0.200 (CAPM). For the FF, Carhart, and 4-index models, the trusts with the medium level of disagreement have a weaker relation between the average premium and lag NAV performance compared to the trusts with the smallest disagreement.

The results in Table 6 suggest that although the coefficients on the disagreement interaction variables are in the main not statistically significant, the magnitude of the coefficients are economically significant. The results suggest that investor heterogeneity has an impact on the relation between the average trust premium and the lag NAV performance. Trusts with the largest disagreement about performance have a weaker relation between the average premium and the lag NAV performance. Our findings support the results in Ferson and Lin(2010), although not as statistically significant, that investor heterogeneity is an important issue in closed-end fund performance studies.

V Conclusions

We use the approach of Ferson and Lin(2010) to examine the impact of investor heterogeneity on the performance of U.K. investment trusts relative to a number of linear factor models. There are three main findings in our study. First, we find no evidence that investment trusts either add or destroy value for investors. There are no trusts with significant superior or inferior performance beyond what we expect by chance. This result differs from Kosowski, Timmerman, Wermers and White(2006) and Fama and French(2010) for U.S. open-end mutual

funds and for Cuthbertson, Nitzsche and O'Sullivan(2008) for U.K. open-end mutual funds where the worst funds have significant inferior performance using net fund returns.

Second, we find that investor heterogeneity has a significant impact on the evaluation of investment trust performance. The magnitude of the bounds in Ferson and Lin(2010) is large in economic terms and often exceeds that of the standard errors of the estimated performance and the impact of alternative benchmarks on performance. We find using unconditional moments, that for the multifactor models nearly all trusts have upper bounds that are significantly different from zero. We find no evidence that investors will agree on the sign of performance for any trust beyond what we would expect by chance. This result, coupled with our first finding, suggests that there are no trusts where all investors can agree that the trust adds value. Using a model of conditional moments, the impact of investor heterogeneity becomes greater and the magnitude of the bounds increase and we can reject the null hypothesis of zero upper bounds for nearly all trusts across all the factor models. These results are similar to Ferson and Lin but the impact is larger here since the upper bound of Ferson and Lin is greater when the funds have higher residual volatility, which we observe in our sample of closed-end funds. The importance of investor heterogeneity on trust performance is also consistent with the evidence in Ahn et al(2009) and Fletcher and Marshall(2012).

Third, we find using the lower bounds of Ferson and Lin(2010) that investor heterogeneity has an impact on the relation between the trust premium and the past NAV performance. We find that trusts with the largest disagreement about performance have a weaker relation between the average premium and past NAV performance. This result is similar to Ferson and Lin who find for open-end U.S. mutual funds that funds with larger disagreement

about performance have a significantly weaker relation between fund flows and past performance.

Our results suggest that investor heterogeneity is an important issue in the evaluation of investment trust performance. Our sample of funds only includes U.K. equity funds. An interesting extension of this study would be to explore the impact of heterogeneity on the performance of international funds. Ferson and Lin(2010) point out that a key issue for future research is the development of clientele specific fund performance measure measures. The NA performance bounds of Ahn et al(2009) provide one way to do this for the clienteles most favorable and least favorable to the fund. However this captures the extremes and a fuller examination of this topic warrants more research. We leave these issues to future research.

Table 1 Summary Statistics of Investment Trusts and Factors

Panel A	Mean	Std Dev	Min	Max
Trusts				
Bottom 1%	-7.654	3.169	-86.921	6.112
Bottom 10%	-0.774	5.121	-60.392	13.470
Bottom 25%	-0.021	5.851	-32.094	17.329
Median	0.365	7.143	-22.867	23.409
Top 25%	0.645	11.062	-17.506	39.431
Top 10%	1.050	18.491	-12.088	85.401
Top 1%	5.842	79.357	-5.897	454.673
Mean	0.248	11.073	-28.772	47.148
Panel B	Mean	Std Dev	Min	Max
Factors				
Market	0.297	4.239	-13.797	10.534
SMB	-0.075	3.815	-12.501	14.535
HML	0.430	2.385	-14.038	7.907
WML	1.015	3.969	-28.403	12.282
Large	0.314	3.928	-14.393	10.667
S-L	-0.010	3.557	-13.555	11.345
AHML	0.268	3.065	-9.649	8.993
S-M	-0.046	2.170	-6.336	7.228
M-L	0.036	2.994	-14.001	11.386
LHML	0.222	3.793	-14.821	12.888
MHML	0.206	5.122	-28.018	17.668
SHML	0.488	3.419	-28.403	17.599

The table reports summary statistics of the monthly excess stock returns of U.K. investment trusts with U.K. equity objectives and factors in the linear factor models between January 1990 and December 2010. The summary statistics include the mean, standard deviation (Std Dev), minimum (Min), and maximum (Max) of monthly excess returns (%). In panel A of the table, all trusts with greater than 24 stock return observations are sorted separately by the statistic in the column and various percentiles of the cross-sectional distribution across individual trusts are reported. Panel B reports the summary statistics for the factors in the linear factor models.

Table 2 Performance of Individual Investment Trusts

Panel A: CAPM	a_p	$SE(a_p)$	$\sigma(u_{pt})$	$a_p(VW)$
Bottom 1%	-8.059	0.096	2.038	-0.008
Bottom 10%	-0.879	0.208	3.375	(-0.06)
Bottom 25%	-0.181	0.329	4.376	
Median	0.095	0.554	5.981	
Top 25%	0.355	1.079	9.763	
Top 10%	0.859	1.973	17.345	
Top 1%	7.632	10.697	76.289	
Mean	0.063	1.057	9.839	
Panel B: FF	a_p	$SE(a_p)$	$\sigma(u_{pt})$	$a_p(VW)$
Bottom 1%	-11.817	0.110	2.023	-0.059
Bottom 10%	-1.228	0.193	3.164	(-0.61)
Bottom 25%	-0.305	0.264	3.749	
Median	-0.007	0.488	5.188	
Top 25%	0.277	0.990	8.933	
Top 10%	0.754	2.187	17.119	
Top 1%	5.887	10.392	70.308	
Mean	-0.228	1.006	9.152	
Panel C: Carhart	a_p	$SE(a_p)$	$\sigma(u_{pt})$	$a_p(VW)$
Bottom 1%	-11.852	0.131	2.007	-0.007
Bottom 10%	-1.120	0.224	3.131	(-0.07)
Bottom 25%	-0.306	0.291	3.743	
Median	0.029	0.531	5.098	
Top 25%	0.372	1.035	8.888	
Top 10%	0.754	2.201	16.916	
Top 1%	6.716	10.562	70.246	
Mean	-0.162	1.059	9.044	
Panel D: 4-index	a_p	$SE(a_p)$	$\sigma(u_{pt})$	$a_p(VW)$
Bottom 1%	-9.680	0.142	2.098	0.079
Bottom 10%	-0.995	0.207	3.154	(0.75)
Bottom 25%	-0.196	0.302	3.705	
Median	0.187	0.525	5.142	
Top 25%	0.528	1.012	8.951	
Top 10%	1.037	2.124	17.112	
Top 1%	5.906	9.864	73.063	
Mean	0.066	1.038	9.101	
Panel E: 7-index	a_p	$SE(a_p)$	$\sigma(u_{pt})$	$a_p(VW)$
Bottom 1%	-13.217	0.141	2.070	0.066
Bottom 10%	-0.994	0.208	3.051	(0.67)
Bottom 25%	-0.245	0.305	3.601	
Median	0.154	0.525	4.978	
Top 25%	0.436	1.044	8.795	
Top 10%	0.939	2.063	16.635	
Top 1%	17.923	15.283	66.851	
Mean	0.114	1.155	8.749	

The table reports percentiles of the cross-sectional distribution of investment trust performance between January 1990 and December 2010. The unconditional Jensen(1968) performance of all trusts with greater than 24 stock return observations is estimated relative to five linear factor models. The percentiles are reported for the Jensen performance (a_p), the standard error of performance ($SE(a_p)$), and the volatility of the regression residuals ($\sigma(u_{pt})$) where the observations are sorted separately by the statistic in the column. The final column reports the performance and t -statistic in parentheses of the value weighted portfolio of all trusts ($a_p(VW)$). The standard errors and t -statistics are corrected for the effects of heteroskedasticity and serial correlation using the automatic lag selection approach (without prewhitening) of Newey and West(1994).

Table 3 Percentiles of $t(a_p)$ Estimates of Simulated and Actual Trust Excess Returns

	CAPM	FF	Carhart	4-index	7-index
Bottom 1%	-2.591 (-3.756) [81.390]	-2.315 (-3.480) [93.100]	-2.649 (-3.879) [84.240]	-2.705 (-4.146) [86.440]	-2.471 (-4.975) [94.910]
Bottom 10%	-1.005 (-1.393) [87.730]	-1.421 (-1.431) [49.810]	-1.403 (-1.480) [59.130]	-1.229 (-1.492) [81.840]	-1.201 (-1.527) [87.060]
Bottom 25%	-0.382 (-0.709) [84.510]	-0.705 (-0.727) [52.950]	-0.666 (-0.748) [62.640]	-0.395 (-0.754) [92.970]	-0.522 (-0.771) [84.820]
Median	0.209 (-0.003) [75.260]	-0.020 (0.001) [46.520]	0.035 (0.001) [56.290]	0.288 (-0.006) [90.230]	0.304 (-0.002) [91.210]
Top 25%	0.743 (0.684) [59.070]	0.633 (0.713) [38.030]	0.717 (0.733) [48.860]	1.132 (0.724) [95.060]	1.026 (0.753) [87.280]
Top 10%	1.112 (1.305) [28.590]	0.978 (1.363) [7.360]	1.265 (1.418) [30.200]	1.615 (1.405) [79.000]	1.462 (1.463) [51.660]
Top 1%	1.752 (2.555) [3.680]	2.330 (2.719) [24.670]	2.667 (2.935) [37.240]	2.708 (2.969) [38.340]	2.544 (3.049) [20.470]

The table reports percentiles of the cross-sectional distribution of actual $t(a_p)$ statistics of individual trust performance between January 1990 and December 2010. In parentheses below are the average simulated (across 10,000 trials) $t(a_p)$ statistic in the () brackets under the null hypothesis of zero trust performance and the proportion (%) of simulated $t(a_p)$ statistics that lie below the corresponding actual $t(a_p)$ statistic across the 10,000 trials in the [] brackets. All trusts are required to have greater than 24 stock return (simulated benchmark-adjusted stock return) observations when computing performance. The performance of the trusts is evaluated relative to five linear factor models. The t -statistics are corrected for the effects of heteroskedasticity and serial correlation using the automatic lag selection approach (without prewhitening) of Newey and West(1994).

Table 4 Upper Bounds on the Impact of Investor Heterogeneity: Unconditional Moments

Panel A	CAPM	FF	Carhart	4-index	7-index
Bottom 1%	0.143	0.345	0.807	0.617	0.694
Bottom 10%	0.237	0.540	1.258	0.927	1.022
Bottom 25%	0.307	0.640	1.504	1.089	1.206
Median	0.420	0.886	2.048	1.511	1.667
Top 25%	0.686	1.525	3.571	2.631	2.946
Top 10%	1.219	2.924	6.796	5.031	5.572
Top 1%	5.362	12.008	28.221	21.481	22.395
Mean	0.691	1.563	3.633	2.675	2.931
SR	0.070	0.171	0.402	0.294	0.335
Panel B	CAPM	FF	Carhart	4-index	7-index
Bottom 1%	0.978	1.461	1.905	1.695	1.826
Bottom 10%	1.264	2.155	3.646	2.851	3.203
Bottom 25%	1.291	2.291	4.314	3.196	3.675
Median	1.315	2.367	4.774	3.372	3.974
Top 25%	1.330	2.412	5.069	3.471	4.110
Top 10%	1.345	2.466	5.270	3.563	4.223
Top 1%	1.409	2.548	5.588	3.783	4.571
N	0	0.984	0.995	0.995	0.995
Panel C	CAPM	FF	Carhart	4-index	7-index
Bottom 1%	-1.223 [56.500]	-7.842 [35.550]	-22.143 [25.610]	-16.935 [23.660]	-9.521 [73.190]
Bottom 10%	-0.426 [48.250]	-1.736 [49.410]	-4.777 [47.250]	-3.362 [48.540]	-3.181 [68.570]
Bottom 25%	-0.261 [44.500]	-0.879 [53.290]	-2.564 [46.230]	-1.709 [53.080]	-1.977 [50.080]
Median	-0.104 [39.430]	-0.554 [51.890]	-1.652 [41.510]	-1.022 [57.850]	-1.178 [59.610]
Top 25%	0.117 [31.170]	-0.341 [45.720]	-1.103 [51.250]	-0.653 [61.760]	-0.845 [53.710]
Top 10%	0.910 [44.970]	0.134 [55.620]	-0.761 [43.150]	-0.323 [54.550]	-0.481 [41.630]
Top 1%	7.327 [53.510]	7.964 [87.900]	2.286 [60.690]	3.382 [56.100]	5.387 [59.650]

The table examines the impact of the upper bound of investor heterogeneity of Ferson and Lin(2010), using unconditional moments, on individual investment trust performance between January 1990 and December 2010. Panel A of the table reports the percentiles of the upper bounds for individual trusts. The final row of panel A reports the adjusted maximum Sharpe ratio (SR) for each set of factor excess returns adjusted for finite sample bias as in Ferson and Siegel(2003) when the number of factors in the benchmark model is greater than 1. Panel B of the table reports percentiles of the estimated t -statistics of the upper bound where the standard errors are estimated as the standard deviation of the simulated upper bounds across 10,000 trials. N is the proportion of trusts where the individual t -statistic is significant at the 5% level using a one-tail test. Panel C of the table reports the percentiles of the actual values of $|a_p|$ - Upper Bound. In parentheses below is the proportion (%) of the simulated $|a_p|$ - Upper Bound which lie below the actual $|a_p|$ - Upper Bound across the 10,000 trials. All trusts are required to have greater than 24 stock return (simulated benchmark-adjusted stock return) observations when computing performance and the upper bound. The performance of the trusts is evaluated relative to five linear factor models.

Table 5 Upper Bounds on the Impact of Investor Heterogeneity: Conditional Moments

Panel A	CAPM	FF	Carhart	4-index	7-index
Bottom 1%	0.466	0.907	1.239	1.032	1.220
Bottom 10%	0.742	1.379	1.858	1.509	1.758
Bottom 25%	0.986	1.675	2.310	1.816	2.078
Median	1.295	2.291	3.111	2.481	2.867
Top 25%	2.101	3.816	5.219	4.247	4.983
Top 10%	3.681	7.111	9.541	7.512	8.700
Top 1%	16.842	29.958	41.131	33.678	36.123
Mean	2.157	3.936	5.363	4.292	4.921
SR	0.229	0.450	0.621	0.494	0.591
Panel B	CAPM	FF	Carhart	4-index	7-index
Bottom 1%	1.595	1.879	2.017	1.884	2.012
Bottom 10%	2.886	3.615	4.116	3.775	3.676
Bottom 25%	3.252	4.308	5.144	4.791	4.698
Median	3.468	4.897	6.148	5.619	5.877
Top 25%	3.597	5.349	6.962	6.297	6.681
Top 10%	3.647	5.529	7.393	6.646	7.056
Top 1%	3.749	5.763	7.727	6.889	7.350
N	0.989	0.995	0.995	0.995	0.995
Panel C	CAPM	FF	Carhart	4-index	7-index
Bottom 1%	-9.271	-25.301	-35.054	-29.133	-19.121
	[53.160]	[29.440]	[29.090]	[28.770]	[77.230]
Bottom 10%	-2.538	-5.566	-7.800	-5.867	-5.988
	[55.540]	[52.880]	[44.940]	[64.500]	[86.170]
Bottom 25%	-1.407	-2.929	-4.049	-3.151	-3.775
	[61.690]	[64.110]	[62.340]	[72.310]	[64.580]
Median	-0.977	-1.890	-2.719	-1.964	-2.234
	[54.280]	[64.430]	[49.930]	[78.940]	[78.310]
Top 25%	-0.637	-1.317	-1.853	-1.369	-1.704
	[50.830]	[70.400]	[67.590]	[85.130]	[78.980]
Top 10%	-0.351	-1.020	-1.526	-1.066	-1.329
	[40.960]	[57.930]	[37.140]	[63.130]	[46.480]
Top 1%	2.838	2.887	-0.109	0.598	0.499
	[59.910]	[91.940]	[31.360]	[37.370]	[19.500]

The table examines the impact of the upper bound of investor heterogeneity of Ferson and Lin(2010), using a model of conditional moments, on individual investment trust performance between January 1990 and December 2010. Panel A of the table reports the percentiles of the upper bounds for individual trusts. The final row of panel A reports the adjusted average conditional maximum Sharpe ratio (SR) for each set of factor excess returns adjusted for finite sample bias as in Ferson and Siegel(2003) when the number of factors in the benchmark model is greater than 1. Panel B of the table reports percentiles of the estimated t -statistics of the upper bound where the standard errors are estimated as the standard deviation of the simulated upper bounds across 10,000 trials. Panel C of the table reports the percentiles of the actual values of $|a_p|$ - Upper Bound. In parentheses below is the proportion (%) of the simulated $|a_p|$ - Upper Bound which lie below the actual $|a_p|$ - Upper Bound across the 10,000 trials. All trusts are required to have greater than 24 stock return (simulated benchmark-adjusted stock return) observations when computing performance and the upper bound. The performance of the trusts is evaluated relative to five linear factor models.

Table 6 Impact of Investor Heterogeneity on Relation between Trust Premium and Past NAV Performance

Panel A	CAPM	FF	Carhart	4-index	7-index
Constant	0.238 (1.26)	0.141 (0.71)	-0.057 (-0.19)	-0.066 (-0.25)	0.136 (0.71)
a _p	3.146 (0.61)	1.031 (0.20)	7.844 (2.41) ¹	3.928 (1.45)	12.271 (2.14) ¹
Total NAV _p	0.034 (2.42) ¹	0.039 (2.84) ¹	0.057 (2.84) ¹	0.056 (3.60) ¹	0.045 (3.33) ¹
DY _p	0.941 (2.41) ¹	1.043 (2.69) ¹	1.410 (3.91) ¹	1.054 (7.09) ¹	1.416 (3.85) ¹
Age _p	0.005 (0.26)	0.005 (0.29)	-0.009 (-0.71)	-0.004 (-0.31)	-0.011 (-0.80)
R ²	0.168	0.164	0.119	0.122	0.137
Panel B	CAPM	FF	Carhart	4-index	7-index
Constant	0.302 (1.55)	0.122 (0.56)	0.104 (1.11)	-0.016 (-0.06)	0.306 (1.97) ¹
a _p	20.073 (2.57) ¹	14.832 (2.07) ¹	10.951 (2.47) ¹	12.072 (2.24) ¹	16.028 (1.89) ²
a _p *Disagree _p	-5.301e+003 (-1.77) ²	-3.473e+003 (-1.34)	-2.295e+003 (-1.08)	-6.703e+003 (-1.64) ²	-6.529e+003 (-1.38)
Total NAV _p	0.029 (2.08) ¹	0.042 (2.89) ¹	0.047 (5.58) ¹	0.052 (3.01) ¹	0.034 (3.01) ¹
DY _p	1.131 (2.73) ¹	0.932 (2.39) ¹	1.250 (8.29) ¹	0.674 (2.40) ¹	1.037 (5.38) ¹
Age _p	0.006 (0.32)	0.004 (0.28)	-0.009 (-0.64)	-0.001 (-0.08)	-0.006 (-0.44)
R ²	0.186	0.188	0.153	0.153	0.170
Panel C	CAPM	FF	Carhart	4-index	7-index
Constant	0.374 (1.81) ²	0.097 (0.41)	-0.012 (-0.07)	-0.061 (-0.19)	0.255 (1.65) ²
a _p	7.127 (1.28)	13.673 (1.21)	14.045 (1.93) ²	11.237 (1.26)	8.685 (1.67) ²
a _p *Disagree _{pH}	-13.984 (-1.35)	-19.843 (-1.12)	-6.719 (-1.13)	-23.468 (-1.23)	-1.317 (-0.13)
a _p *Disagree _{pM}	12.890 (1.08)	-11.842 (-0.80)	-8.114 (-1.30)	-9.007 (-0.77)	6.776 (0.63)
Total NAV _p	0.024 (1.60)	0.043 (2.56) ¹	0.054 (5.07) ¹	0.056 (2.81) ¹	0.034 (3.14) ¹
DY _p	1.065 (2.43) ¹	0.510 (0.95)	1.396 (5.07) ¹	0.556 (1.47)	1.249 (5.60) ¹
Age _p	0.008 (0.42)	0.006 (0.35)	-0.008 (-0.61)	-0.005 (-0.34)	0.001 (0.07)
R ²	0.201	0.175	0.129	0.155	0.168

¹ Significant at 5%

² Significant at 10%

The table reports the results of Fama and MacBeth(1973) cross-sectional regressions of the annual average monthly premium during the calendar year on a constant, lagged performance (a_p) using the Net Asset Value (NAV) excess returns of the trusts, and the interaction between the lagged performance and the measure of the disagreement among investors about trust performance ($Disagree_p$). The $Disagree_p$ measure is estimated using the NAV excess returns of the trusts. The regressions are run each year between 1993 and 2010. The performance of the trusts is estimated during the prior 36 months relative to each factor model for all trusts with continuous NAV return data during the past 3 years and the $Disagree_p$ measure is assumed fixed across the whole sample period. The cross-sectional regressions include control variables such as the total NAV of the trust, the dividend yield (DY), and the age of the trust at the start of the calendar year. Each panel reports the time-series average coefficients and t -statistics in parentheses. The final row of each panel is the time-series average of the adjusted R^2 . In panel A of the table, the cross-sectional regressions are run excluding the interaction $Disagree_p$ variables. In panel B of the table, the actual value of the disagreement measure is used and in panel C of the table, two dummy variables are used for the disagreement measure. The $Disagree_{pH}$ variable is 1 for trusts with the largest 1/3 of $Disagree_p$ measures and 0 otherwise, and the $Disagree_{pM}$ dummy variable is 1 for trusts with the second largest 1/3 of $Disagree_p$ measures and 0 otherwise. The t -statistics are corrected for the effects of heteroskedasticity and serial correlation using the automatic lag selection approach (without prewhitening) of Newey and West(1994).

References

- Ahn, D.H., Cao, H.H. and S. Chretien, 2009, Portfolio performance measurement: A no arbitrage bounds approach, *European Financial Management*, 15, 298-339.
- Aragon, G. and W.E. Ferson, 2008, Portfolio performance evaluation, *Foundations and Trends in Finance*, 2, 83-190.
- Bal, Y. and L.A. Leger, 1996, The performance of U.K. investment trusts, *Services Industry Journal*, 16, 67-81.
- Bangassa, K., 1999, Performance of U.K. investment trusts 1980-1994, *Journal of Business Finance and Accounting*, 26, 1141-1168.
- Berk, J.B. and R. Stanton, 2007, Managerial ability, compensation, and the closed-end fund discount, *Journal of Finance*, 62, 529-556.
- Brown, S. J., Goetzmann, W.N., Ibbotson, R. and S.A. Ross, 1992, Survivorship bias in performance studies, *Review of Financial Studies*, 5, 553-580.
- Carhart, M. M., 1997. Persistence in mutual fund performance. *Journal of Finance*, 52, 57-82.
- Carhart, M.M., Carpenter, J.N., Lynch, A.W. and D.K. Musto, 2002, Mutual fund survivorship, *Review of Financial Studies*, 15, 1439-1463.
- Chen, Z. and P.J. Knez, 1996, Portfolio performance measurement: Theory and applications, *Review of Financial Studies* 9, 511-555.
- Cherkes, M., Sagi, J. and R. Stanton, 2009, A liquidity-based theory of closed-end funds, *Review of Financial Studies*, 22, 257-297.
- Cochrane, J.H., 2005, *Asset pricing: Revised edition* (Princeton NJ: Princeton University Press).
- Cochrane, J.H. and J. Saa-Requejo, 2000, Beyond arbitrage: Good-deal asset pricing bounds in incomplete markets, *Journal of Political Economy*, 108, 79-119.

Connor, G. and R.A. Korajczyk, 1986, Performance measurement with the arbitrage pricing theory: A new framework for analysis, *Journal of Financial Economics*, 15, 373-394.

Cremers, M., Petajisto, A. and E. Zitzewitz, 2012, Should benchmark indices have alpha? Revisiting performance evaluation, *Critical Finance Review*, forthcoming.

Cuthbertson, K., Nitzsche, D. and N. O'Sullivan, 2008, Mutual fund performance: Skill or luck?, *Journal of Empirical Finance*, 15, 613-634.

Da, Z. and H. Yun, 2010, Electricity consumption and asset prices, *Working Paper*, University of Notre Dame.

Dimson, E. and C. Minio-Paluello, 2001, The closed-end fund discount and performance persistence, *Working Paper*, London Business School.

Dimson, E. and C. Minio-Paluello, 2002, The closed-end fund discount, *Research Monograph*, CFA Institute.

Elton, E., Gruber, M.J. and J. Busse, 1988, Do investors care about sentiment?, *Journal of Business*, 71, 477-500.

Fama, E.F. and K.R. French, 1993, Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33, 3-56.

Fama, E.F. and K.R. French, 2010, Luck versus skill in the cross section of mutual fund returns, *Journal of Finance*, 65, 1915-1948.

Fama, E.F., and J.D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy*, 71, 607-636.

Ferson, W.E., 2010, Investment performance evaluation, *Annual Review of Financial Economics*, 2, 207-234.

Ferson, W.E., 2012, Investment performance: A review and synthesis, In Constantinides, G.M., Harris, M., and R. Stultz (eds) Handbook of the Economics of Finance, Elsevier Science Publishers, North Holland, forthcoming.

Ferson, W.E. and J. Lin, 2010, Alpha and performance measurement: The effect of investor heterogeneity, *Working Paper*, University of Southern California.

Ferson, W.E. and A.F. Siegel, 2003, Stochastic discount factor bounds with conditioning information, *Review of Financial Studies*, 16, 567-595.

Ferson, W.E. and A.F. Siegel, 2009, Testing portfolio efficiency with conditioning information, *Review of Financial Studies*, 22, 2735-2758.

Fletcher, J. and D. Forbes, 2002, U.K. unit trust performance: Does it matter which benchmark or measure is used, *Journal of Financial Services Research*, 21, 195-218.

Fletcher, J. and A.P. Marshall, 2012, Evaluating U.K. investment trust performance using no arbitrage bounds, *Working Paper*, University of Strathclyde.

Griffin, J.M., 2002, Are the Fama and French factors global or country-specific?, *Review of Financial Studies*, 15, 783-803.

Grinblatt, M. and S. Titman, 1994, A study of monthly mutual fund returns and performance evaluation techniques, *Journal of Financial and Quantitative Analysis*, 29, 419-444.

Hansen, L.P. and R. Jagannathan, 1991, Implications of security market data for models of dynamic economies, *Journal of Political Economy*, 99, 225-262.

Hansen, L.P. and S.F.R. Richard, 1987, The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models, *Econometrica*, 55, 587-613.

Harrison, M and D. Kreps, 1979, Martingales and arbitrage in multi-period securities markets, *Journal of Economic Theory*, 20, 381-408.

Jensen, M.C., 1968, The performance of mutual funds in the period 1945-1964, *Journal of Finance* 23, 389-416.

Khorana, A.J., Servaes, H. and L. Wedge, 2009, Portfolio manager ownership and the pricing of closed-end funds, *Working Paper*, London Business School.

Kosowski, R., Timmerman, A., Wermers, R. and H. White, 2006, Can mutual fund “stars” really pick stocks? New evidence from a bootstrap analysis, *Journal of Finance*, 61, 2551-2596.

Lehmann, B.N. and D.M. Modest, 1987, Mutual fund performance evaluation: A comparison of benchmarks and benchmark comparisons, *Journal of Finance*, 42, 233-265.

Lettau, M. and S.C. Ludvigson, 2010, Measuring and modeling variation in the risk-return tradeoff, In Ait-Sahalia, Y. and L.P. Hansen (eds) *Handbook of Financial Econometrics*, Elsevier Science Publishers, North Holland, 618-690.

Linnainmaa, J.T., 2012, Reverse survivorship bias, *Journal of Finance*, forthcoming.

Malkiel, B.G., 1977, The valuation of closed-end investment-company shares, *Journal of Finance*, 32, 847-858.

Newey, W.K. and K.D. West, 1994, Automatic lag selection in covariance matrix estimation, *Review of Economic Studies*, 61, 631-653.

Ramadorai, T., 2012, The secondary market for hedge funds and the closed hedge fund premium, *Journal of Finance*, 67, 479-512.

Ross, S.A., 1978, A simple approach to the valuation of risky streams, *Journal of Business*, 51, 153-475.