Evaluating Mutual Fund's Alpha via Alternative Frameworks: Some New Evidence and Insights

Sharon Garyn-Tal* and Beni Lauterbach**

This version: November 1, 2011

Abstract

This paper documents comprehensive and updated evidence on the performance of and correlation between various regression frameworks for estimating "alpha" - the excess return offered by the fund to its investors. Examining over 1000 U.S non-specialized mutual funds in 2001-2009, our main findings are: 1) alpha's magnitude, the fund's classification as good or poor, and the fund's ranking relative to other funds, all strongly depend on the regression framework. The differences are substantial even among somewhat-similar-specification regression models; 2) funds' performance persistence can be traced. However, persistence is strongly dependent on the evaluation period, and sometimes evaporates once more complex (and probably more correct) pricing-models are used as benchmarks; 3) when we compute the net economic alpha (an alpha that takes into account the fact that the alternative of investing in ETFs is also costly), the average net economic alpha across the 2001-2009 period is below 1% per year in absolute value, regardless of the regression framework used. Thus, the net economic alpha correction, employed for the first time in this study, portrays the mean return of the mutual fund industry as only slightly inferior to ETFs. In addition, given that the net economic alpha measure provides a more accurate estimate of the excess return offered to fund investors, it should receive more attention.

* Faculty of Industrial Engineering and Management, Technion - Israel Institute of Technology, Haifa 32000, Israel. E-mail: sharon.tal.garyn@gmail.com.

** School of Business Administration, Bar-Ilan University, Ramat-Gan 52900, Israel.
Email: lauteb@mail.biu.ac.il.

Acknowledgements: We thank Simon Benninga, Doron Kliger, Haim Reisman and Avi Wohl for their constructive comments. All remaining errors are our own.

1 Introduction

The classical measures of investment performance compare the return of a managed portfolio to the return of a benchmark portfolio. The benchmark portfolio should represent an equivalent and a feasible investment alternative to the managed portfolio being evaluated.

The Capital Asset Pricing Model (Sharpe ,1964) implies that all investors should hold a broadly diversified portfolio - the market portfolio - and safe assets in a portion according to their tastes for risk. Jensen (1968) uses the market portfolio of the CAPM to advocate the Jensen's alpha. A positive alpha implies the manager earns an abnormal return relative to the alternative of holding the benchmark portfolio. Other measures of portfolio performance developed by the early literature on portfolio performance are the Sharpe ratio and the Treynor's measure (see Treynor (1965)). In general, this strand casts doubt on the ability of portfolio managers to outperform passive indices.

Following the CAPM, the Arbitrage Pricing model of Ross (1976) allows for several risk factors to determine assets' expected returns, but leaves it up to empirical research to identify the risk factors. Current practice follows Fama and French (1996) and Carhart (1997), who suggest evaluating performance using three- or four-factor models derived from empirically observed patterns in stock portfolio returns. Further elaborations include adding a style factor, represented by the average return of other managed portfolios in the same market sector (see Hunter et al (2009)).

Previous research suggest that alternative risk-adjusted benchmarks lead to substantive differences in performance measures. For example, Lehmann and Modest (1987) find that funds' rankings are very sensitive to the asset-pricing model chosen to measure performance. Elton et al. (1993) also stress the importance of choosing the correct benchmark. They expand the model used in Ippolito (1989) and show that Ippolito's results can be reversed when using different benchmarks.[1]

[1] Roll (1978), Grinblatt and Titman (1994), Chan, Dimmock and Lakonishok (2006), and Coles, Daniel and Nardari (2006) also show that estimated performance is sensitive to the choice of the reference portfolio.

This paper compares different benchmark models and evaluates their influence on estimating the fund manager's contribution, alpha, to the return of her investors. We estimate the alphas via several benchmark models, some of which are novel or only recently developed. Using a sample of nine years (2001-2009) and over 1000 non-specialized open-end equity funds' returns, we offer the most comprehensive comparison to date of benchmark impact on fund performance. Not surprising, alpha's magnitude depends on the benchmark model applied for funds' performance evaluation. More surprising, there exist substantial differences in funds rankings and classifications even for somewhat similar specifications of the benchmark models.

In existing comparisons of a managed portfolio return with that of a performance benchmark it is regularly assumed that the benchmark has no costs. This generates a bias against the evaluated managed portfolio as the measured return of the managed portfolio, mutual fund for example, is evaluated net of the funds' expenses and trading costs. Fama and French (2010) write that "...if the question is whether mutual funds are better for investors than passive investments, benchmark returns, like fund returns, should be net of costs."

We compute the net economic alpha by adding to the traditional alpha as extracted from a benchmark model the cost of mimicking the fund's systematic risk via ETFs. The motivation for the net economic alpha is that the alternative to buying a mutual fund - buying ETFs on the benchmarks - is not costless (for example - the costs of mimicking multiple benchmarks or the long-short portfolios of Fama and French results is non-negligible). Our corrected net economic alpha should better represent the net economic (i.e, considering alternatives) excess return of the fund. This study is the first to compute and discuss net economic alphas. Interestingly, we also show that the net economic alpha can be approximated by the alpha extracted from a regression of the net of fee fund return on the net of fee benchmark return.

Comparing the standard alphas with their corresponding net economic alphas, we find little differences in fund's classification and ranking. In our data, the impact of the net economic alpha is minute. However, interestingly, the mean net economic alpha of mutual funds in our sample is typically small, between 0 and -1% per year. Moreover, examining the subperiods of our sample, the mean net economic alphas are becoming more positive with time; In the last subperiod net economic alphas are all positive. This may imply that the mutual fund industry is only slightly inferior to ETFs and that it (the mutual fund industry) is becoming more and more competitive (relative to ETFs) with time.

Another substantial body of research examines the persistence of mutual fund returns. Persistence is a crucial issue for both investors and fund managers[2]. Previous research documents mutual fund return predictability and attributes it to managers' different information, stock-picking talent, "hot hands" or common investment strategies[3]. At the same time, there is also contrary evidence of performance reversion[4].

We test funds' performance persistence implied by the different benchmark models. We define a first evaluation period of three years and compare it with a later performance period of three years. The results imply that persistence is strongly dependent on the evaluation period and on the evaluation method, and sometimes evaporates once more complex (and probably more correct) pricing-models are used as benchmarks. The CAPM yields inconsistent persistence results with statistically significant positive persistence at the first period and statistically significant negative persistence at the second period.

[2]For example, see Ippolito (1989), Brown et al. (1996), Gruber (1996), Chevalier and Ellison (1997), Sirri and Tufano (1998), Zheng (1999), Busse (2001), Goriaev et al. (2005) and others.

[3]For example, see Sharpe (1966), Treynor(1965)), Carlson (1970), Ippolito (1992), Grinblatt and Titman (1992), Brown, Goetzmann, Ibbotson and Ross (1992), Hendricks, Patel and Zeckhauser (1993), Jegadeesh and Titman (1993), Goetzmann and Ibbotson (1994), Shukla and Trzcinka (1994), Hendricks et al. (1993, 1997), Brown and Goetzmann (1995), Malkiel (1995) Grinblatt et al. (1995), Gruber (1996), Elton et al. (1996), Carhart (1997), Graham and Harvey (1997), Otten and Bams (2002), Berk and Green (2004), Wermers (2004), Bollen and Busse (2005), Kaplan and Schoar (2005), Kosowski, Timmermann, Wermers and White (2006), Cremers and Petajisto (2009) and many more. On the other hand, Chen, Jegadeesh and Wermers (2000), for example, find weak evidence that funds with the best past performance have better stock-picking skills than funds with the worst past performance.

[4]For example, Carhart (1997) finds very slight evidence consistent with skilled or informed mutual fund managers. Winners are somewhat more likely to remain winners, and losers are more likely to either remain losers or perish. Furthermore, last year's winners frequently become next year's losers and vice versa, which is consistent with the gambling behavior of mutual funds. Thus, while the ranking of the few of the top and many at the bottom funds persist, the year-to-year ranking on most funds appear largely random.

We conclude that one needs to specify her risk or benchmark model before selecting a fund. In addition, for the results to be more objective (and comparable to ETF investment), the net economic alpha correction, as suggested in this paper, should be adopted.

2 Net Economic Alphas - Definition and Measurement

Consider a managed portfolio P. We attempt to evaluate the fund manager contribution to the return of her investors, alpha, by applying different frameworks for evaluating alpha. We also suggest (and later test) an elaboration of existing methods.

First lets assume that the fund is operating within the CAPM framework and extract alpha via the standard regression of the excess return of the portfolio P $(R_P - R_f)$ on the market excess return $(R_M - R_f)$. Traditionally, this alpha is used as a standard measure for the fund manager performance. However, a slight modification appears necessary. Note that the alternative to buying a mutual fund - buying an ETF on the market - is not costless. Thus, perhaps a more relevant measure of performance should add to the traditional alpha, as extracted from the regression, the cost of mimicking the fund's systematic risk via an ETF.[5] This is the net economic alpha version of the traditional performance measure.

A second method employs as a benchmark the fund's style. The fund's style is determined by the benchmark it follows. It is represented by the return on the index of the style (factor S). We extract alpha from the regression of the fund's excess return on the style factor's excess return. Then, to assess the net economic alpha, we add back to alpha the cost of mimicking the fund's risk, i.e, the fund's exposure to its style (β_S) times the cost of investing in an ETF that follows the fund's style index.

[5]Fama and French (2010) write that "...if the question is whether mutual funds are better for investors than passive investments, benchmark returns, like fund returns, should be net of costs." Thus, Fama and French advocate the use of the net economic alpha as discussed and tested for the first time in the present study.

A third possible benchmark is a similar fund, represented by the average return of funds belonging to the fund's group (factor G). The funds can be gathered into groups based on fund's style. Recall that the fund's style is determined by the benchmark it follows; Thus, all funds that follow the same benchmark constitute a group. Funds can also be gathered into groups based on fund's classification. We run a regression of the fund's excess return on the group (factor G) average excess return, and estimate an alpha from that regression. There are some serious limitations to factor G. Most notably, the mean alpha is forced (by definition) to equal zero (see Appendix 1). Hence, we can only infer fund's relative performance.

It can also be argued that a fund is operating within a multi-factor APT model. We consider five multi-factor models: a three- and a four-factor models (Fama French and Carhart), and three five-factor models, based on Hunter et al (2009), adding the group factor or the style factor to the Carhart's four-factor alpha estimation. In these multiple factor frameworks (other than those that include the group-factor) we also attempt to add back to the regression-computed alpha the costs of mimicking the fund's systematic risk in order to estimate the net economic excess return of the fund.

2.1 Net Economic Alphas in APT Frameworks - A Formal Derivation

Assume, for simplicity, an APT model with one factor - the market (factor M). According to the APT model

$$E(R_P) - R_f = \beta_{P,M}(E(R_M) - R_f). \tag{1}$$

Let the realized excess return of portfolio P be

$$\tilde{R}_P - \tilde{R}_f = \beta_{P,M}(\tilde{R}_M - \tilde{R}_f) + \tilde{\alpha}_{P,M} + \tilde{\epsilon}_{P,M}. \tag{2}$$

We define the last term in the regression ($\tilde{\alpha}_{P,M} + \tilde{\epsilon}_{P,M}$) as the return on the portable alpha[6] portfolio, portfolio X. Portfolio X is a zero investment portfolio that has zero covariance with the benchmark. In equilibrium, $\alpha_{P,M}$ - the expected return of X -

[6] Portable alpha strategies have been discussed by Usman and Gold (2001), Lee (2005), Kung and Pohlman (2005), Anson (2005), Ezrati (2006), Gorman and Weigand (2007), Hubric (2008) and others.

should be zero. However, in practice, $\alpha_{P,M}$ is a result of fund manager's activity and talent - a measure of performance.

The portfolio P return, equation (2), can be represented in the following way

$$\tilde{R}_P = \beta_{P,M}\tilde{R}_M + (1 - \beta_{P,M})\tilde{R}_f + \tilde{R}_X. \qquad (3)$$

In equation (3) we decompose portfolio P return into a convex combination of: 1. the market index, 2. a money market component, and 3. a portfolio X having zero covariance with the benchmark index.

The only elaboration we offer to the basic approach is the computation of the net economic alpha - the alpha that takes into account the cost of the alternative of investing in ETFs.

Suppose that the expense ratio for the benchmark M, defined as the cost of investing in an ETF mimicking the market index M, is a percentage ϕ_M of assets, and that the expense ratio for the fund P is a percentage ϕ_P of assets. Then, the cost of mimicking the benchmarked part of the managed portfolio P appears to be $\phi_M\beta_{P,M}$. However, at this point we must reserve and offer a correction. Note that $\phi_M\beta_{P,M}$ implies that when beta is negative the fee for the benchmarked part of the managed portfolio P is negative. This contradicts the reality that a short position, taken by the fund manager relative to factor M (a negative $\beta_{P,M}$) costs at least as much as a long position in M. Thus, if beta with respect to the market, factor M, is negative, the cost of replicating this systematic (benchmarked) part of the portfolio should be positive, and should be estimated as: $|\beta_{P,M}|\phi_M$. Therefore, the fee charged by the fund on the active (non-systematic) part of the portfolio equals $\phi_P - \phi_M|\beta_{P,M}|$.[7]

Recall that in the standard regressions for estimating alpha we use the net of fee excess returns (based on the net of fee prices of the mutual fund) as the dependent variable. This implies that the standard methodology of regressing $\tilde{R}_P - \tilde{R}_f$ on $\tilde{R}_M - \tilde{R}_f$ estimates the net of fee alpha ($\alpha_{P,M}$). Hence, the gross (before fees) excess return of the fund (raw alpha) is $\alpha_{P,M} + \phi_P$.

[7]We assume negligible costs for the money market component of the portfolio.

If we deduct from the raw alpha the fee charged on the active part, we obtain $\alpha_{NET,P,M}$, the net economic alpha:

$$\alpha_{NET,P,M} = \alpha_{P,M} + \phi_P - (\phi_P - \phi_M|\beta_{P,M}|).$$

Thus,

$$\alpha_{NET,P,M} = \alpha_{P,M} + \phi_M|\beta_{P,M}|. \tag{4}$$

Since $\alpha_{NET,P,M}$ takes into account the cost of mimicking the fund's risk, it better represents the fund manager's contribution to her investors.

Similarly, when we consider an APT model with N factors, then

$$\tilde{R}_P = \beta_{P,0} + \sum_{i=1}^{N} \beta_{P,i}\tilde{R}_i + \tilde{\epsilon}_P. \tag{5}$$

and an analogous equation to (4) can be derived:

$$\alpha_{NET,P,N} = \alpha_{P,N} + \sum_{i=1}^{N} \phi_i|\beta_{P,i}|, \tag{6}$$

where $\alpha_{P,N}$ is the intercept of the regression of the (net of fee) excess return of the fund on the excess returns of the factors in the multi-factor model, ϕ_i is the expense ratio for an index or an ETF that follows benchmark i, and $\beta_{P,i}$ is benchmark i's coefficient extracted from the regression of the (net of fee) excess return of the fund on the excess returns of the factors in the multi-factor model. $\alpha_{NET,P}$ is exactly the same modified performance measure that our intuition suggested.

Another interesting and intuitively appealing point is that $\alpha_{NET,P}$, the net economic alpha, can be approximated by the intercept in a regression of net of fees mutual fund excess return on net of fees factors excess returns. A net of fee factor excess return is the excess return on an ETF that mimics the factor minus the fees of such an ETF. More details on this point and a formal derivation are offered in Appendix 2.

3 Benchmark Models

Academic researchers use different benchmark-frameworks for evaluating funds performance. The benchmark-framework typically represents a feasible investment strategy, comparable to the fund portfolio.

The Capital Asset Pricing Model (Sharpe,1964) suggests as a benchmark the market return and it is the first and traditional performance evaluation methodology.[8] Chan, Dimmock and Lakonishok (2006) and Cremers and Petajisto (2009) also use a one factor model, yet replace the market return with the return on an index reflecting fund's style.[9]

The Arbitrage Pricing model of Ross (1976) allows for several risk factors to determine assets' expected returns. Academic practice follows Fama and French (1996) and Carhart (1997), who evaluate mutual funds using three- or four-factor models (see also Otten and Bams (2002), Wermers (2004), Fama and French (2010)). The three factors are: 1. the excess market returns, 2. the performance of value stocks relative to growth stocks, and 3. the performance of small stocks relative to big stocks. The fourth factor is the momentum factor (MOM) of Carhart (1997), constructed as the average return on the two high prior return portfolios minus the average return on the two low prior return portfolios.[10]

Another school of performance evaluation employs the average return on similar mutual funds as a benchmark. We call it the group factor. This methodology uses the average return of a group of mutual funds as a benchmark for each individual

[8]Gruber (1996), Carhart (1997), Sirri and Tufano (1998), Otten and Bams (2002) use the CAPM for evaluating mutual funds performance.

[9]Researchers use different methodologies for identifying the style benchmark of a fund. For example, Chan, Dimmock and Lakonishok (2006) use a portfolio's weighted average size percentile rank across its holdings to determines the manager's size orientation, and a manager's composite value score to determines the value/growth orientation. On the other hand, Cremers and Petajisto (2009) compute the Active Share of a fund with respect to all the indexes they consider and assign the index with the lowest Active Share as that fund's benchmark.

[10]Quoting from French web site: "a momentum factor, constructed from six value-weight portfolios formed using independent sorts on size and prior return of NYSE, AMEX, and NASDAQ stocks. UMD is the average of the returns on two (big and small) high prior return portfolios minus the average of the returns on two low prior return portfolios. The portfolios are constructed daily. Big means a firm is above the median market cap on the NYSE at the end of the previous day; small firms are below the median NYSE market cap. Prior return is measured from day - 250 to - 21. Firms in the low prior return portfolio are below the 30th NYSE percentile. Those in the high portfolio are above the 70th NYSE percentile."

fund within that group. The group benchmarking adjusts for commonalities in similar actively managed mutual fund returns. Similar funds should have similar loadings on common priced factors (including factors omitted from popular models), and similar choices of specific stocks and industries.

In its most developed form, Hunter, Kandel, Kandel and Wermers (2009) use a group factor based on the average return of all funds belonging to the fund's classification group. Hunter et al also examine a five-factor model that includes the four factors of Carhart (1997) plus the group factor, where, again, the group factor is defined according to the fund's classification.

We also examine the group factor and the five-factor Hunter et al (2009) model. In addition, we propose and examine a variation of the group-benchmark methodology. This variation defines a fund's group according to the index with which its return is most correlated. Sometimes, the official fund classification does not accurately represent its behavior. Thus, actual style benchmarking, as proposed by Chan, Dimmock and Lakonishok (2006) and Cremers and Petajisto (2009), might be more appropriate. Practically, we advance two models that, to the best of our knowledge, have not been examined before: 1) a group factor model with the group factor defined as the average return of all funds belonging to the same style group as the fund, 2) a five-factor model that includes the four factors of Carhart (1997) and the style-based group factor defined above.

Last, we also propose a five-factor model comprising the four factors of Carhart (1997) plus a fifth benchmark - the return on the index that represents the fund's style. The model is also, to the best of our knowledge, novel.

4 Data and Benchmark Models

4.1 Sample and Variables

We collect monthly data on mutual fund returns in 2001-2009 from CRSP (The Center for Research in Security Prices). The CRSP data uses the Lipper classification that divides the world of non-specialized open-end equity funds into 12 groups based on funds' style: LCCE (Large-Cap Core Funds), LCGE (Large-Cap Growth Funds), LCVE (Large-Cap Value Funds), MCCE (Mid-Cap Core Funds), MCGE (Mid-Cap Growth Funds), MCVE (Mid-Cap Value Funds), SCCE (Small-Cap Core Funds), SCGE (Small-Cap Growth Funds), SCVE (Small-Cap Value Funds), MLCE (Multi-Cap Core Funds), MLGE (Multi-Cap Growth Funds), and MLVE (Multi-Cap Value Funds). We prefer the Lipper classification because it divides the universe of non-specialized open-end equity funds into 12 categories, more than other classification we considered. However, the Lipper classification of our funds starts only in 2001; thus our data starts in 2001.

For the empirical analysis we need data on several other variables. The Fama-French-Carhart four factors returns are taken from French website:
http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.[11]

We plan to define a fund's style by identifying the benchmark that yields the highest correlation with respect to the fund's return. Thus, a list of benchmarks should be suggested. Based on a review of the internet information sheet of Fidelity and Vanguard funds that match the 12 funds' objectives above. We find that the main benchmarks used for evaluating these funds' performance are the S&P500 or an appropriate Russell index (Russell 1000, Russell 1000 growth, Russell 1000 value, Russell midcap, Russell midcap growth, Russell midcap value, Russell 2000, Russell 2000 growth, Russell 2000 value). The Russell indices data and returns are extracted from the Russell investments website (www.russell.com). S&P500 returns are collected from the CRSP database.

[11]The four Fama-French factors are: 1. the performance of small stocks relative to big stocks (SMB, Small Minus Big), 2. the performance of value stocks relative to growth stocks (HML, High Minus Low), based upon the Fama-French Portfolios, and 3. the excess market returns, based on the value-weighted return on all NYSE, AMEX, and NASDAQ stocks (from CRSP), and the one-month Treasury bill rate (from Ibbotson Associates). The momentum factor is the average return on the two high prior return portfolios minus the average return on the two low prior return portfolios; Quoting from French web site: "a momentum factor, constructed from six value-weight portfolios formed using independent sorts on size and prior return of NYSE, AMEX, and NASDAQ stocks. UMD is the average of the returns on two (big and small) high prior return portfolios minus the average of the returns on two low prior return portfolios. The portfolios are constructed daily. Big means a firm is above the median market cap on the NYSE at the end of the previous day; small firms are below the median NYSE market cap. Prior return is measured from day - 250 to - 21. Firms in the low prior return portfolio are below the 30th NYSE percentile. Those in the high portfolio are above the 70th NYSE percentile."

4.2 Variables Constructed For Net Economic Alpha Estimation

Since we are interested in net economic alphas, we need to assess the expense ratios of ETFs replicating fund styles, and expense ratios of ETFs replicating the Fama-French-Carhart four factors.

We collect data mainly from two large ETFs issuers: iSHARES, the largest issuer of long position ETFs, and ProShares, the largest issuer of short position ETFs.

Assessing the expense ratios of ETFs, we ignore brokerage commissions because all kinds of discount and on-line trading are available, and we omit the bid-ask spread because it is negligible for heavily traded ETFs. As for mutual funds expenses, we also ignore redemption fees and loads. We hope that these ignored costs of ETFs and mutual funds balance each other off.

Based on a review of the internet information sheet of ETFs that follow the 10 indices specified above and additional ETFs that match the benchmarks examined in the study (see Table in Appendix 3), we assess that a long (short) position in the S&P500 - the market factor - costs 0.09% (0.92%), a long (short) position in Russell 1000 costs 0.15% (0.95%), a long (short) position in Russell 1000 growth costs 0.2% (0.95%), a long (short) position in Russell 1000 value costs 0.2% (0.95%), a long (short) position in Russell midcap costs 0.21% (0.95%), a long (short) position in Russell midcap growth costs 0.25% (0.95%), a long (short) position in Russell midcap value costs 0.26% (0.95%), a long (short) position in Russell 2000 costs 0.26% (0.95%), a long (short) position in Russell 2000 growth costs 0.25% (0.95%), and a long (short) position in Russell 2000 value costs 0.4% (0.95%).

The expense ratio of the momentum factor is composed of the expense ratio charged for investing in a long position in a winners portfolio (via PowerShares DWA Technical Leaders Portfolio that charges an expense ratio of 0.6%) as well as of the expense ratio charged for investing in a short position in a losers portfolio (via JETS Contrarian Opportunities Index Fund that charges 0.58&). Thus, we assess that a long (short) position in the momentum factor costs 1.18% (1.18%).

The expense ratio of a long position in the SMB factor is composed of the expense ratio charged for investing in a long position in a small stocks portfolio (via iShares

Russell 2000 Index Fund that charged an expanse ratio of 0.26%) plus the expense ratio charged for investing in a short position in a big stocks portfolio (via Direxion Daily Large Cap Bear 3x Shares that charges an expense ratio of 0.95%). Thus, we assess that a long position on the SMB factor costs 1.21%. A short position in the SMB factor is composed of a short position in the small stocks portfolio (via ProShares Short Russell2000 that charges an expense ratio of 0.95%) and of a long position in the big stocks portfolio (via iShares Russell 1000 Index Fund that charged an expense ratio of 0.15%). Thus, we assess that a short position in the SMB factor costs 1.1%.

In the same spirit, we assess that a long (short) position in the HML factor costs 1.2% (1.2%).

4.3 The Benchmark Models

We apply nine different benchmark models for evaluating a fund's alpha. The models differ in the benchmark-system used in the regression.

1. M - the CAPM regression is

$$R_{P,t} - R_{f,t} = \alpha_P + \beta_P (R_{M,t} - R_{f,t}) + \epsilon_{P,t}.$$

The excess market return is based on the value-weight return on all NYSE, AMEX, and NASDAQ stocks (from CRSP) and on the one-month Treasury bill rate (from Ibbotson Associates).

2. S - the fund's style - the style benchmark is determined by checking the correlation between a fund's return and the return on the following ten indices - Russell 1000, Russell 1000 growth, Russell 1000 value, Russell midcap, Russell midcap growth, Russell midcap value, Russell 2000, Russell 2000 growth, Russell 2000 value, and S&P500 - and choosing the index that yields the highest correlation with respect to the fund's return. The regression equation for fund P is

$$R_{P,t} - R_{f,t} = \alpha_P + \beta_P (R_{S,t} - R_{f,t}) + \epsilon_{P,t}.$$

3. G(Class) - the average return of all funds belonging to the fund's (Lipper) classification group. The regression equation for fund P is

$$R_{P,t} - R_{f,t} = \alpha_P + \beta_P(R_{G(Class),t} - R_{f,t}) + \epsilon_{P,t}.$$

4. G(Style) - the average return of all funds belonging to the fund's style group. The regression equation for fund P is

$$R_{P,t} - R_{f,t} = \alpha_P + \beta_P(R_{G(Style),t} - R_{f,t}) + \epsilon_{P,t}.$$

5. 3F - the Fama and French three-factor model is

$$R_{P,t} - R_{f,t} = \alpha_P + \beta_{P,1}(R_{M,t} - R_{f,t}) + \beta_{P,2}HML_t + \beta_{P,3}SMB_t + \epsilon_{P,t}.$$

The factors are: 1. the excess market returns is based on the value-weight return on all NYSE, AMEX, and NASDAQ stocks (from CRSP) and on the one-month Treasury bill rate (from Ibbotson Associates), 2. the performance of value stocks relative to growth stocks (HML, High Minus Low), and 3. the performance of small stocks relative to big stocks (SMB, Small Minus Big), based upon the Fama-French Portfolios.

6. 4F - Carhart's (1997) four-factor model is

$$R_{P,t} - R_{f,t} = \alpha_P + \beta_{P,1}(R_{M,t} - R_{f,t}) + \beta_{P,2}HML_t + \beta_{P,3}SMB_t + \beta_{P,4}MOM_t + \epsilon_{P,t}.$$

The model is based on the Fama and French three-factor model and the additional momentum factor (MOM) of Carhart (1997), constructed as the average return on the two high prior return portfolios minus the average return on the two low prior return portfolios.[12].

[12]Quoting from French web site: "a momentum factor, constructed from six value-weight portfolios formed using independent sorts on size and prior return of NYSE, AMEX, and NASDAQ stocks. UMD is the average of the returns on two (big and small) high prior return portfolios minus the average of the returns on two low prior return portfolios. The portfolios are constructed daily. Big means a firm is above the median market cap on the NYSE at the end of the previous day; small firms are below the median NYSE market cap. Prior return is measured from day - 250 to - 21. Firms in the low prior return portfolio are below the 30th NYSE percentile. Those in the high portfolio are above the 70th NYSE percentile."

7. 4F-G(Class) - Hunter et al (2009) five-factor model, where the group factor is defined according to fund's Lipper classification.[13] The regression equation for fund P is

$$R_{P,t} - R_{f,t} = \alpha_P + \beta_{P,1}(R_{M,t} - R_{f,t}) + \beta_{P,2}HML_t + \beta_{P,3}SMB_t + \beta_{P,4}MOM_t + \beta_{P,5}R_{G^*(Class),t} + \epsilon_{P,t}.$$

8. 4F-G(Style) - a five-factor model, where the group factor is defined according to the fund's style group.[14] The regression equation for fund P is

$$R_{P,t} - R_{f,t} = \alpha_P + \beta_{P,1}(R_{M,t} - R_{f,t}) + \beta_{P,2}HML_t + \beta_{P,3}SMB_t + \beta_{P,4}MOM_t + \beta_{P,5}R_{G^*(Style),t} + \epsilon_{P,t}.$$

9. 4F-S - a five-factor model, where the fifth benchmark is the fund's style.[15] The regression equation for fund P is

$$R_{P,t} - R_{f,t} = \alpha_P + \beta_{P,1}(R_{M,t} - R_{f,t}) + \beta_{P,2}HML_t + \beta_{P,3}SMB_t + \beta_{P,4}MOM_t + \beta_{P,5}R_{S^*,t} + \epsilon_{P,t}.$$

Hereafter, we refer to models 7-8 as the Hunter et al five-factor models, and to models 7-9 as the five-factor models.

4.4 Data Characteristics

The data comprises nine years (2001-2009) of funds' returns and is divided into three subperiods: 2001-2003, 2004-2006, and 2007-2009.

Panel A of Table 1 reports the 2001-2003 data characteristics. There are 1037 funds existing throughout 2001-2003, of which 188 are classified as LCCE, 123 are classified as LCGE, 66 are classified as LCVE, 52 are classified as MCCE, 85 are classified as MCGE, 34 are classified as MCVE, 68 are classified as MLCE, 68 are classified as MLGE, 98 are classified as MLVE, 89 are classified as SCCE, 100 are

[13] Hunter et al (2009) regress the average group return on the other benchmarks return, and use the intercept plus residual of this regression as an estimate of factor G*.

[14] We regress the average group return on the other benchmarks return, and use the intercept plus residual of this regression as an estimate of factor G*.

[15] We regress the style factor return on the other benchmarks return, and use the intercept plus residual of this regression as an estimate of factor S*.

classified as SCGE, and 66 are classified as SCVE. The 2001-2003 average Total Net Assets (TNA) of all funds is 667.9 million dollars. The average expense ratio (management fee) charged by fund managers is 1.31% (0.74%) and the weighted (by TNA) average expense ratio (management fee) is 1.24% (0.74% as well).[16] Khorana, Servaes and Tufano (2009) study fees charged by mutual funds in 2002. They document (in Table 2, page 1288) an average management fee of 0.62% per year and an average expense ratio of 1.11% per year for U.S equity funds. These fees are slightly lower than the fees we report above.

Panel B reports the 2004-2006 data characteristics. There are 1547 funds existing throughout 2004-2006, with at least 64 funds in each of the 12 Lipper groups. The 2004-2006 average TNA of all funds is 971.2 million dollars. The average expense ratio (management fee) charged by fund managers is 1.37% (0.78%) and the weighted (by TNA) average expense ratio (management fee) is 1.28% (0.78% as well).

Panel C reports the 2007-2009 data characteristics. There are 1312 funds existing throughout 2007-2009, with at least 48 funds in each of the Lipper groups. The 2007-2009 average TNA of all funds is 759.6 million dollars. The average expense ratio (management fee) charged by fund managers is 1.27% (0.73%) and the weighted (by TNA) average expense ratio (management fee) is 1.12% (0.73% as well).

(Insert Table 1 about here)

Finally, since we distinguish in our empirical work between fund's (Lipper) classification and fund's style, where style is the benchmark most correlated with fund's return, we examine the relation between Lipper classification and (our assessment of) fund's style. Panel D of Table 1 reports, per each Lipper classification, the average distribution of the fund's style. The results are not surprising. For example, in the Large-Cap Growth funds group (LCGE raw in Panel D) 72.4% of the funds correlate most the Russell 100 growth index, a benchmark that we found most common in these type of funds. Nevertheless, these findings highlight the fact that Lipper classification groups are not homogeneous. Thus, we decide to construct

[16] Mutual funds' expense ratio includes management fee and 12b1 fee and might also include other expenses. It does not include loads or transaction fees.

groups of funds based on style (i.e, on the index most correlated with fund's return). Such style groups are perhaps more homogeneous, hence it is interesting to find what results they yield.

5 Empirical Results

5.1 Funds' Alpha Estimates

We fit our benchmark models over three periods: 2001-2003, 2004-2006 and 2007-2009. In each period we evaluate all funds with returns throughout all 36 months of that period. On average, we examine 1,299 funds per period. Table 2 reports regression results for 2001-2003 (Panel A), 2004-2006 (Panel B), 2007-2009 (Panel C) and average results over these three subperiods (Panel D). Each Panel summarizes the regression results for each of the nine benchmark frameworks.

We examine first the average results. As reported in Panel D, the mean R-square of our benchmark models varies between 0.85 and 0.94; the minimum mean R-square is obtained using the CAPM regression and the maximum mean R-square is obtained using the five-factor models.

(Insert Table 2 about here)

The mean alphas vary between -1.73% for the three-factor model and -0.46% for the CAPM. Consistent with previous evidence (see Sharpe (1966), Jensen (1968), Grinblatt and Titman (1989), Carhart (1997)), mutual funds tend to deliver disappointing net returns. The negative excess returns can be explained, on average, by the funds' expense ratios, suggesting that the before-fees average performance of the funds is fair and adequate.

Next, we take into consideration the costs of mimicking the funds' risk and construct the net economic alpha. This is approximately the same as computing alpha by a regression of the net of fee fund return on the net of fee benchmarks' returns. The mean net economic alphas, computed for the first time by us, vary between -0.93% for the three-factor model and -0.01% for the four-factor plus the style-factor model. As expected, the mean net economic alphas are closer to zero, portraying mutual funds in a more positive way.

The differences between the alpha and net economic alpha columns in Panel D vary with the complexity of the benchmark model. The three- to five-factor benchmark models naturally entail higher mimicking costs than the one-factor portfolio benchmark models. Thus, for the multi-factor models, the correction for expenses is of higher magnitude, and is non-negligible. The more complicated the model gets, the less trivial is mimicking fund's risk with the various factors, the higher are the alternative mimicking costs (the costs of mimicking fund's risk via ETFs) and the wider are the discrepancies between raw alphas and net economic alphas.

The average alpha of benchmark models that include the group factor equals zero by definition (see Appendix 2 for a proof). For that reason, the average alpha and net economic alpha of models that include the group factor are denoted as N.R (Not Relevant). The group-factor methodology allows only within a group classification and ranking.

An interesting observation is that the CAPM yields the highest alpha (positive and statistically significant) in 2001-2003 (Panel A) and 2007-2009 (Panel C) and the lowest alpha (negative and statistically significant) in 2004-2006 (Panel B). In contrast, all our other models yield, in all three periods, negative alphas that are mostly statistically significant. Evidently, the CAPM alphas are most (and perhaps too) volatile.

Using a t-test of differences, we find that the mean alpha is statistically significant different from the mean net economic alpha at the 1% significance level for all nine benchmark-model and throughout the three periods: 2001-2003, 2004-2006 and 2007-2009. Yet, the magnitude of the mean difference between alphas and net economic alphas is quite modest - less than 1% per year. This implies that economically the difference between alpha and net economic alpha is small and of secondary importance.

Examining the mean net economic alphas of 2001-2009, they are typically between 0 and -1% per year in our sample. The mean net economic alpha is negative and statistically significant for all benchmark models other than the CAPM in 2001-2003 (the CAPM assess a positive and statistically significant positive net economic alpha). In the 2004-2006 subperiod, the mean net economic alpha is negative according to the one factor models and positive according to the three- to five-factor models. In the last subperiod of 2007-2009, the mean net economic alphas are positive for all benchmark models. These findings suggests that on average the mutual fund industry is only slightly inferior to ETFs. More importantly, given the positive alphas in the last subperiod, it appears that the mutual fund industry is becoming more and more competitive (relative to ETFs) with time.

Table 3 reports the correlations between the alphas. The correlations between the alphas vary from 0.48, between the CAPM and the 4F-G(Class), to 0.98, between the three- and four-factor models. Thus, there is a wide agreement between the benchmark models. We use a t-test and find that all correlation coefficients are different from 0 and different from 1 at the 1% significance level.

(Insert Table 3 about here)

5.2 Differences in Funds Classification and Ranking

First, each fund is classified by each benchmark model as either a good fund (positive alpha) or a poor fund (negative alpha). Then, we compare funds classifications. If benchmark model i and benchmark model j agree on fund classification, then there is no classification difference between models i and j. On the other hand, if a fund is classified as good (poor) fund according to benchmark model i and is classified as poor (good) fund according to benchmark model j, then there is a classification difference between models i and j. Table 4 summarizes the results.

(Insert Table 4 about here)

As reported in Panel A, the frequency of fund classification differences varies between from 6% and 34.1%. Minor classification disagreements exist between the three-factor and the four-factor models (6%). Substantial classification differences emerge when comparing the CAPM and the three-, four- and five-factor models (28.1%-34.1%). Other one-factor models yield also substantial classification differences relative to the three- to five-factor models.

Next, we test the null hypothesis that there is no relation between the funds' classification according to the different benchmark models. If classification is independent across models, all average frequencies of classification differences, reported in Panel A, should equal 50%. We find that in Panel A all classification difference frequencies are lower than 0.5 at the 1% significance level. Thus, there is a positive dependence in funds' classification across our nine benchmark models - most of the time our benchmark models tend to agree on a fund's classification.

Panel B presents ranking differences. Assume N funds are available. For each benchmark model, the best performing fund with the highest alpha is ranked at the 1st place, the second best performing fund with the second highest alpha is ranked at the 2nd place, and so on. Since each period a fund has nine different alphas, all funds are ranked nine times. Assume fund P is ranked in the k-th place according to benchmark model i (i.e, according to $\alpha_{i,P}$), and assume that the fund is ranked in the l-th place according to benchmark model j (i.e, according to $\alpha_{j,P}$). If $|k-l| < 0.1N$, i.e the ranking difference is less than 10% of the N existing funds, then we, somewhat arbitrarily, denote that there is no ranking difference between models i and j. On the other hand, If $|k-l| \geq 0.1N$, then there is a ranking difference between models i and j. Panel B reports the 2001-2003, 2004-2006 and 2007-2009 average frequency of ranking differences between the benchmark models.

The ranking-difference frequency varies between 12.9% to 64.5%. Small ranking differences emerge when comparing the three- and four-factor models (12.9%). Substantial ranking differences exist when comparing all other models, especially when comparing the CAPM with all other models (56.5%-64.5%).

For Panels C and D analysis, we narrow the data sample and keep only funds ranked as 20 best performing funds or as 20 worst performing funds. Each period, each benchmark model has its own list of funds included in the 20 best and in the 20 worst performing funds. Then, we count how many funds appear on both benchmark method i and benchmark method j lists of best (worst) performing funds, and report the numbers in Panel C (D).

In Panel C, the average ranking overlap of the 20 best performing funds varies between 5 and 16 funds. The highest overlap is obtained when comparing the three- and four-factor models. The lowest overlap of 5-6 out of 20 best performing funds is obtained when comparing the CAPM with the three versions of the five-factor models.

In Panel D, the average ranking overlap of the 20 worst performing funds vary from 8.7 to 17.3. The highest overlap emerges when comparing the three- and four-factor models. The lowest overlap occurs when comparing the style-factor model and the Hunter et al five-factor model (in which the group factor is based on funds' classification).

In sum, the evidence in this subsection, supports the view that a fund's classification and ranking depends significantly on the benchmark-system used for its evaluation. Differences in fund's classification and ranking are substantial even for somewhat similar specifications of the benchmark models. Notably, the CAPM benchmark appears most extreme - it generates the highest disagreement proportions in Table 4. Basing a performance evaluation solely on the CAPM benchmark appears erroneous.

5.3 Differences Due To Net Economic Alpha Calculation

In each period - 2001-2003, 2004-2006 and 2007-2009, we calculate net economic alphas of our sample funds, and compare them with the corresponding standard alphas, for each of the nine benchmark models. Basically, for each raw alpha analyzed in the previous subsection, we compute a corresponding net economic alpha (see section 2 for definition).

Table 5 reports the comparison results. Panel A compares funds classification differences (a positive versus a negative alpha), Panel B reports 10% and more differences in the funds rankings, and Panel C (D) compares the 20 best (worst) performing funds.

(Insert Table 5 about here)

Panel A reports, for each benchmark method, the 2001-2003, 2004-2006 and 2007-2009 average frequency of classification differences between alpha and net economic alpha. Classification difference are rare for the CAPM (0.8%) and relatively infrequent for the four-factor plus the style-factor model (12.7%). In the three, four and five-factor models, the alternative of mimicking a fund's risk with ETFs is more costly than in one-factor models. Hence, naturally, fund classification differences increase with benchmark model complexity. Generally, the emerging differences in funds' classification in Panel A appear small. Only seldom does the net economic alpha methodology change the fund classification suggested by raw alphas.

Panel B reports the 2001-2003, 2004-2006 and 2007-2009 average frequency of ranking differences between the alphas and the net economic alphas. As reported in Panel B, no ranking differences exist for the one-factor models, and small ranking differences exist for the three, four and five-factor models (2.3%-4%).

We also generate the lists of 20 "best" funds and 20 "worst" funds according to net economic alphas, and compare them with the corresponding lists generated based on raw alphas. The list of 20 best and 20 worst performing funds mostly overlap. The overlap is 17.7-20 out of 20 funds for the best performing funds and 17.3-20 for the worst performing funds.

In summary, the evidence in this section suggests that, in our data, alphas and net economic alphas yield almost identical inference, even when multi-factor models and models that include pricey factors such as the long-short portfolios are employed. This can also be interpreted as encouraging because it suggests that previous research, based solely on raw alphas, is probably correct, i.e, similar results would be obtained with net economic alphas. Nevertheless, since net economic alphas are the more "decent" and "fair" measures to judge fund's performance, future studies should examine them as well. One should also recall that in Table 2, the net economic alpha analysis suggests that the mean underperformance of mutual funds (relative to ETFs) is slight and that the mutual fund industry relative performance improved over time. Thus, the net economic alphas, computed for the first time in our paper, afford some important economic inference.

5.4 Performance Persistence Tests

Next, we test funds' performance persistence implied by the different benchmark models. We define two periods - the first (period 1) is a three-year "evaluation" period and the second (period 2) is a three-year "performance" period. We use all available funds and check for the correlation between period 1 and period 2 following performance measures: 1. the funds' alphas, 2. the funds' ranks (based on funds' alphas), where the ranks of the funds at the performance period include all funds existing at that period, whether they had existed at the evaluation period or not (original ranks), 3. the funds' ranks (based on funds' alphas), where the ranks of the funds at the performance period include only surviving funds from the evaluation period (veteran surviving ranks), and 4. funds' deciles belongings (based on funds' alphas).

Panel A of Table 6 reports the correlations between funds' alpha during the 2001-2003 evaluation period and the 2004-2006 performance period. The data comprises 728 funds existing from 2001 till 2006. Panel B of Table 6 reports the correlations between funds' alpha during the 2004-2006 evaluation period and the 2007-2009 performance period. The data comprises 866 funds existing from 2004 till 2009. The correlations are calculated and reported for each of the nine benchmark models.

The correlations implied by the four types of performance measures (alpha, original ranks, veteran surviving ranks, and decile ranks) are of similar magnitude. For each of the nine benchmark models, we next report the average (across the four performance measures) correlation.

(Insert Table 6 about here)

In Panel A, the between-periods performance correlation varies from -0.06 for the four-factor model to 0.33 for the CAPM. The one-factor models imply a higher performance persistence (correlation of 0.14-0.33) than the three- to five-factor models (for which the average correlations vary between -0.06 and 0.06). The correlations implied by the one-factor and two-factor models are different from zero at the 1% significance level, while the correlations implied by the three to five-factor models are mostly insignificant. Evidently, when more complex pricing models are used, performance persistence tends to evaporate. The CAPM yields the highest performance persistence (with a correlation of 0.33), whereas the performance persistence implied by the four-factor model is negative and insignificant.

In Panel B, the between-periods performance correlation of all models other than the CAPM is between 0.11 and 0.23. The between-periods performance correlation of the CAPM is -0.22. All the coefficients are different from zero at the 1% significance level.

We also compute the average persistence across our nine benchmark models and report it on the bottom of Panels A and B. The between-periods alpha correlations are positive on average and statistically significant. However, these correlations are mostly low - about 0.1, suggesting that persistence is weak and perhaps economically insignificant.

6 Summary and Conclusion

In this paper we compare nine different benchmark models, some of which are novel or only recently developed, and evaluate their impact on estimating the fund manager's contribution, alpha, to the return of her investors. We use a sample of over 1000 non-specialized open-end equity funds' returns extracted from CRSP. The sample period is 2001-2009.

We begin by estimating funds' alpha over three-year periods using nine benchmark models. Consistent with previous literature, we find that funds, on average, do not beat their benchmarks; Rather funds underperform the benchmark. Alpha's magnitude depends on the benchmark model used for funds' performance evaluation. There is a wide agreement between the benchmark models - the minimum correlation between funds alpha is 0.48. On the other hand, one might argue that the extent of agreement between the models is disappointing. When we compare the ranking and classification of funds based on the alphas as extracted from the nine benchmark models, significant classification and ranking differences are detected even among "similar" benchmark models, with the traditional CAPM benchmark demonstrating the strongest differences relative to all other methods. We conclude that one needs to specify her risk or benchmark model before selecting a fund.

Next, we compare the alphas and the corresponding net economic alphas. The net economic alpha is computed by adding to the traditional alpha as extracted from a benchmark model the cost of mimicking the fund's systematic risk via ETFs. Theoretically, the net economic alpha should be used for evaluating the real contribution of a fund manager to her investors, since it takes into account the alternative cost of replicating funds' risk (see Fama and French, 2010).

This study is the first to compute net economic alphas. Comparing the standard and the net economic alphas, little differences in funds classification and ranking emerge, even for the multi-factor models and for models that include pricey factors such as the long-short portfolios. Thus, in our data, the impact of the net economic alpha is minute. However, interestingly, the mean net economic alpha of mutual funds in our sample is below 1% per year in absolute value, regardless of the benchmark or pricing model used. Examining the subperiods in our sample, the mean net economic alphas become more positive over time; in the last subperiod all mean net economic alphas are positive. This may imply that the mutual fund industry is only slightly inferior to ETFs and that the mutual fund industry is becoming more and more competitive relative to ETFs with time. This may be the important "macro" implication of our study.

Next, we examine the influence of using the different benchmark models on funds' performance persistence. We define a first evaluation period of three years and a later performance period of three years. We document evidence consistent with funds' performance persistence, as between-periods alpha correlations are positive and statistically significant. These correlations are however low (about 0.1) on average, suggesting that persistence is weak and perhaps economically insignificant.

Interestingly, the CAPM yields inconsistent persistence results, with a statistically significant indication of persistence at the first test period, and a statistically significant indication of lack of persistence (i.e., performance reversal) at the second test period. The four- and five-factor models' between periods alpha correlations are more robust, yet almost always closer to zero in absolute value. It appears that performance persistence tends to evaporate once more complex (and probably more correct) pricing standards are used.

References

P. U. Ali and M. L. Gold. An overview of "portable alpha" strategies, with practical guidance for fiduciaries and some comments on the prudent investor rule. Working Paper, University of Queensland Law School, 2001.

M. Anson. Institutional portfolio management. *Journal of Portfolio Management*, 31(4):33–43, 2005.

J. B. Berk and R. C. Green. Mutual fund flows and performance in rational markets. *Journal of Political Economy*, 112:1269–1295, 2004.

K. Brown, W. Harlow, and L. Starks. Of tournaments and temptations: An analysis of managerial incentives in the mutual fund industry. *Journal of Finance*, 51:85–110, 1996.

S. J. Brown and W. N. Goetzmann. Performance persistence. *Journal of Finance*, 50:679–698, 1995.

S. J. Brown, W. N. Goetzmann, R. G. Ibbotson, and S. A. Ross. Survivorship bias in performance studies. *Review of Financial Studies*, 5(4):553–580, 1992.

J. A. Busse. Another look at mutual fund tournaments. *Journal of Financial and Quantiative Analysis*, 36:53–73, 2001.

M. M. Carhart. On persistence in mutual funds performance. *Journal of Finance*, 52(1):57–82, 1997.

R. S. Carlson. Aggregate performance of mutual funds, 1948- 1967. *Journal of Financial and Quantitative Analysis*, 5:1–32, 1970.

L. K. C. Chan, S. G. Dimmock, and J. Lakonishok. Benchmarking money manager performance: Issues and evidence. Working Paper NBER No. 12461, 2006.

H.-L. Chen, N. Jegadeesh, and R. Wermers. The value of active mutual fund management: An examination of the stockholdings and trades of fund managers. *Journal of Financial and Quantitative Analysis*, 35(3):343–368, 2000.

J. Chevalier and G. Ellison. Risk-taking by mutual funds as a response to incentives. *Journal of Political Economy*, 105(6):1167–1200, 1997.

J. L. Coles, N. D. Daniel, and F. Nardari. Does the choice of model or benchmark affect inference in measuring mutual fund performance? Arizona State University Working Paper, 2006.

M. Cremers and A. Petajisto. How active is your fund manager? a new measure that predicts performance. *Review of Financial Studies*, 22:3329–3365, 2009.

E. J. Elton, M. J. Gruber, D. Sanjiv, and M. Hlavka. Efficiency with costly information: A reinterpretation of evidence from managed portfolios. *Review of Financial Studies*, 6:1–22, 1993.

E. J. Elton, M. J. Gruber, and C. R. Blake. Survivorship bias and mutual fund performance. *Review of Financial Studies*, 9:1097– 1120, 1996.

M. Ezrati. Portable alpha: A glimpse at the future. *Journal of investing*, 15(3): 119–123, 2006.

E. F. Fama and K. R. French. Multifactor explanations of asset pricing anomalies. *Journal of Finance*, 51:55–84, 1996.

E. F. Fama and K. R. French. Luck versus skill in the cross-section of mutual fund returns. *Journal of Finance.*, 65(5):1915–1947, 2010.

W. N. Goetzmann and R. G. Ibbotson. Do winners repeat? *Journal of Portfolio Management*, 20(2):9–18, 1994.

A. P. Goriaev, T. E. Nijman, and B. J. Werker. Yet another look at mutual fund tournaments. *Journal of Empirical Finance*, 12:127–137, 2005.

L. R. Gorman and R. A. Weigand. Measuring alpha based performance. implications for alpha focused, structured products. Working Paper, 2007.

J. R. Graham and H. R. Campbell. Grading the performance of market-timing newsletters. *Financial Analysts Journal*, 53:54–66, 1997.

M. Grinblatt and S. Titman. Mutual fund performance: An analysis of quarterly portfolio holdings. *Journal of Business*, 62(3):393–416, 1989.

M. Grinblatt and S. Titman. The persistence of mutual fund performance. *Journal of Finance*, 47:1977–1984, 1992.

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

M. Grinblatt, S. Titman, and R. Wermers. Momentum investment strategies, portfolio performance and herding: A study of mutual fund behavior. *American Economic Review*, 85:1088–1105, 1995.

M. J. Gruber. Another puzzle: The growth in actively managed mutual funds. *Journal of Finance*, 51:783–810, 1996.

D. Hendricks, J. Patel, and R. Zeckhauser. Hot hands in mutual funds: Short run persistence of relative performance, 1974-1988. *Journal of Finance*, 48:93–130, 1993.

D. Hendricks, J. Patel, and R. Zeckhauser. The j-shape of performance persistence given survivorship bias. *Review of Economics and Statistics*, 79:161–166, 1997.

S. Hubrich. An alpha unleashed: Optimal derivative portfolios for portable alpha strategies. Working Paper, 2008.

D. Hunter, E. Kandel, S. Kandel, and R. Wermers. Endogenous benchmarks. Working Paper, 2009.

R. A. Ippolito. Efficiency with costly information: A study of mutual fund performance. *Quarterly Journal of Economics*, 104:1–23, 1989.

R. A. Ippolito. Consumer reaction to measures of poor quality: Evidence from the mutual fund industry. *Journal of Law and Economics*, 35:45–70, 1992.

N. Jegadeesh and S. Titman. Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance*, 48:65–91, 1993.

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

S. N. Kaplan and A. Schoar. Private equity returns: Persistence and capital flows. *Journal of Finance*, 60:1791–1823, 2005.

R. Kosowski, A. Timmermann, R. Wermers, and H. White. Can mutual fund stars really pick stocks? new evidence from a bootstrap analysis. *Journal of Finance*, 61:2551–2595, 2006.

E. Kung and L. Pohlman. An alpha + beta framework. *Journal of Investing*, 14(4): 9–16, 2005.

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

B. G. Malkiel. Returns from investing in equity mutual funds 1971 to 1991. *Journal of Finance*, 50:549–572, 1995.

R. Otten and D. Bams. European mutual fund performance. *European Financial Management*, 8(1):75–101, 2002.

R. Roll. Ambiguity when performance is measured by the securities market line. *Journal of Finance*, 33:1051–1069, 1978.

S. A. Ross. The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, 13:341–360, 1976.

W. F. Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19:425–442, 1964.

W. F. Sharpe. Mutual fund performance. *Journal of Business*, 39(1):119–138, 1966.

R. Shukla and C. Trzcinka. Persistent performance in the mutual fund market: Tests with funds and investment advisors. *Review of Quantitative Finance and Accounting*, 4:115–135, 1994.

E. R. Sirri and P. Tufano. Costly search and mutual fund flows. *Journal of Finance*, 53:1589–1622, 1998.

L. R. Thomas. Engineering an alpha engine. *Journal of Investing*, 14(4):23–31, 2005.

J. L. Treynor. How to rate management of investment funds. *Harvard Business Review*, 43:63–75, 1965.

R. Wermers. Is money really smart? new evidence on the relation between mutual fund flows, manager behavior, and performance persistence. Working Paper, University of Maryland, 2004.

L. Zheng. Is money smart? a study of mutual fund investors' fund selection ability. *Journal of Finance*, 54:901–933, 1999.

APPENDIX 1: Limitations of The Group-Factor

A methodological issue occurs whenever using the group factor G in a regression, whether as an only factor or as an additional factor in a multi-factor framework.

First, assume that factor G is the only factor in the regression. Factor G's return is an equally weighted average of all N funds with the same classification C or with the same style S: $\tilde{R}_G = \sum_{i=1}^{N} X_i \tilde{R}_i = \frac{\sum_{i=1}^{N} \tilde{R}_i}{N}$.

For fund i, the regression equation on factor G is

$$\tilde{R}_i - R_f = \alpha_i + \beta_i(\tilde{R}_G - R_f) + \tilde{\epsilon}_i. \tag{7}$$

Summing up the N regression equations of the N funds with the same classification C (or style S) gives

$$\sum_{i=1}^{N} \tilde{R}_i - \sum_{i=1}^{N} R_f = \sum_{i=1}^{N} \alpha_i + \sum_{i=1}^{N} \beta_i(\tilde{R}_G - R_f) + \sum_{i=1}^{N} \tilde{\epsilon}_i.$$

Dividing the equation above by N gives

$$\frac{\sum_{i=1}^{N} \tilde{R}_i}{N} - R_f = \frac{\sum_{i=1}^{N} \alpha_i}{N} + \frac{\sum_{i=1}^{N} \beta_i}{N}(\tilde{R}_G - R_f) + \frac{\sum_{i=1}^{N} \tilde{\epsilon}_i}{N},$$

so

$$\tilde{R}_G - R_f = \bar{\alpha} + \bar{\beta}(\tilde{R}_G - R_f). \tag{8}$$

For fund i, $\beta_i = \frac{cov(\tilde{R}_G, \tilde{R}_i)}{var(\tilde{R}_G)} = \frac{cov(\frac{\sum_{j=1}^{n} \tilde{R}_j}{N}, \tilde{R}_i)}{var(\tilde{R}_G)}$. Thus, the average beta is $\bar{\beta} = \frac{\sum_{i=1}^{N} \beta_i}{N} = \frac{\sum_{i=1}^{N} cov(\frac{\sum_{j=1}^{n} \tilde{R}_j}{N}, \tilde{R}_i)}{N var(\tilde{R}_G)} = \frac{cov(\frac{\sum_{j=1}^{n} \tilde{R}_j}{N}, \frac{\sum_{i=1}^{N} \tilde{R}_i}{N})}{var(\tilde{R}_G)} = \frac{cov(\tilde{R}_G, \tilde{R}_G)}{var(\tilde{R}_G)} = 1$.

Note that in equation (8) if $\bar{\beta} = 1$ then $\bar{\alpha} = 0$. In practice, however, we first run each fund regression separately, and then take the average of the N regressions. Thus, $\bar{\beta}$ might not exactly equal one and $\bar{\alpha}$ might not exactly equal zero.

The exact same point can be shown when factor G is an additional factor in a multi-factor framework. Factor G's return is an equally weighted average of all N funds with the same classification C (or with the same style S): $\tilde{R}_G = \sum_{i=1}^{N} X_i \tilde{R}_i = \frac{\sum_{i=1}^{N} \tilde{R}_i}{N}$. Factor G is calculated following Hunter et al (2009) methodology, and it equals $\alpha_G + \tilde{\epsilon}_G$ as extracted from the regression of factor G return on the other M-1 factors returns (for simplicity, we ignore R_f)

$$\tilde{R}_G = \sum_{j=1}^{M-1} \beta_j \tilde{R}_j + \alpha_G + \tilde{\epsilon}_G. \tag{9}$$

For fund i, the regression framework consists M factors, including factor G,

$$\tilde{R}_i = \alpha_i + \sum_{j=1}^{M-1} \beta_{i,j}\tilde{R}_j + \beta_{i,G}\tilde{R}_G + \tilde{\epsilon}_i.$$

Summing up the N regression equations of all funds with the same classification C (or style S) gives

$$\sum_{i=1}^{N} \tilde{R}_i = \sum_{i=1}^{N} \alpha_i + \sum_{i=1}^{N} \sum_{j=1}^{M-1} \beta_{i,j}\tilde{R}_j + \sum_{i=1}^{N} \beta_{i,G}\tilde{R}_G + \sum_{i=1}^{N} \tilde{\epsilon}_i.$$

Dividing the equation above by N gives

$$\frac{\sum_{i=1}^{N} \tilde{R}_i}{N} = \frac{\sum_{i=1}^{N} \alpha_i}{N} + \frac{\sum_{i=1}^{N} \sum_{j=1}^{M-1} \beta_{i,j}}{N}\tilde{R}_j + \frac{\sum_{i=1}^{N} \beta_{i,G}}{N}\tilde{R}_G + \frac{\sum_{i=1}^{N} \tilde{\epsilon}_i}{N},$$

so

$$\tilde{R}_G = \bar{\alpha} + \sum_{j=1}^{M-1} \bar{\beta}_j\tilde{R}_j + \bar{\beta}_G\tilde{R}_G.$$

From the definition of factor G

$$\tilde{R}_G = \bar{\alpha} + \sum_{j=1}^{M-1} \bar{\beta}_j\tilde{R}_j + \bar{\beta}_G(\alpha_G + \tilde{\epsilon}_G). \tag{10}$$

Combining equations 9 and 10, $\bar{\beta}_G = 1$ and thus $\bar{\alpha} = 0$. Note that in practice $\bar{\beta}_G$ and $\bar{\alpha}$ might not be equal to 1 and 0, respectively, for the same reason explained above.

A problem with $\bar{\alpha} = 0$ is that we cannot tell anymore if one group of funds is superior to another group. Only within a group classification and ranking is possible.

APPENDIX 2: Net Economic Alphas - The Regression Approach

Alpha is traditionally calculated from a regression of net (denote as N) of fees fund excess return ($R_{P,N} - R_f$) on the gross (denote as B) excess return on the appropriate indexes (for index i: $R_{i,B} - R_f$). Thus, $\alpha_{N,B}$ is extracted from the expected value of the regression

$$\bar{R}_{P,N} - R_f = \sum_{i=1}^{N} \beta_{P,N,B,i}(\bar{R}_{i,B} - R_f) + \alpha_{P,N,B}.$$

It could be argued that the basic asset pricing models abstract from imperfections such as transaction costs. Thus, α should be calculated using net returns to the investor. That is, alpha should be calculated in a regression of net excess return of the fund on the net excess return on the appropriate indices ($\alpha_{N,N}$). In this section we will investigate $\alpha_{N,N}$ and its relation to the α_{NET} from section 2.

Taking the expected value of a regression of the net of fee fund return, $R_{P,N}$, on the net of fee benchmarks return, $R_{i,N}, i = 1..N$,

$$\bar{R}_{P,N} - R_f = \sum_{i=1}^{N} \beta_{P,N,N,i}(\bar{R}_{i,N} - R_f) + \alpha_{P,N,N},$$

yields an alpha, noted as $\alpha_{P,N,N}$ (the "net on net" alpha).

Subtracting the traditional alpha from the "net on net" alpha

$$\begin{aligned}
(*)\alpha_{P,N,N} - \alpha_{P,N,B} &= \\
&= [\bar{R}_{P,N} - R_f - \sum_{i=1}^{N} \beta_{P,N,N,i}(\bar{R}_{i,N} - R_f)] \\
&- [\bar{R}_{P,N} - R_f - \sum_{i=1}^{N} \beta_{P,N,B,i}(\bar{R}_{i,B} - R_f)] \\
&= [\sum_{i=1}^{N} \beta_{P,N,B,i}(\bar{R}_{i,B} - R_f)] - [\sum_{i=1}^{N} \beta_{P,N,N,i}(\bar{R}_{i,N} - R_f)] \\
&= \sum_{i=1}^{N} [\beta_{P,N,B,i}(\bar{R}_{i,B} - R_f) - \beta_{P,N,N,i}(\bar{R}_{i,N} - R_f)].
\end{aligned}$$

The relation between $R_{i,N}$ and $R_{i,B}$ depends on whether the beta of fund P with respect for benchmark i ($\beta_{P,i}$) is positive or negative. For a positive beta, i.e fund P takes a long position in benchmark i, the relation between the benchmark's net of fee return and gross of fee return is $\tilde{R}_{i,N} = (1 + \tilde{R}_{i,B})(1 - \phi_i) - 1$. In other words, from the point of view of fund P, when taking a long position in the benchmark, the benchmark's expense ratio (ϕ_i) lowers the benchmark's net of fee return with respect to the benchmark's gross of fee return. The opposite occurs when fund P takes a short position in the benchmark. From the point of view of fund P, when taking a short position in the benchmark, the benchmark's expense ratio (ϕ_i) raises the benchmark's net of fee return with respect to the benchmark's gross of fee return. Thus, for a negative beta, the relation between the benchmark's net of fee return and the benchmark's gross of fee return is $\tilde{R}_{i,N} = (1 + \tilde{R}_{i,B})(1 + \phi_i) - 1$.

First, assume that all betas are positive. Thus, $\tilde{R}_{i,N} = (1+\tilde{R}_{i,B})(1-\phi_i)-1$. For benchmark i, $\beta_{P,N,B,i} = \frac{cov(\tilde{R}_{i,B}-R_f,\tilde{R}_{P,N}-R_f)}{var(\tilde{R}_{i,B}-R_f)}$ and $\beta_{P,N,N,i} = \frac{cov(\tilde{R}_{i,N}-R_f,\tilde{R}_{P,N}-R_f)}{var(\tilde{R}_{i,N}-R_f)}$. Thus,

$$
\begin{aligned}
\beta_{P,N,N,i} &= \frac{cov(\tilde{R}_{i,N} - R_f, \tilde{R}_{P,N} - R_f)}{var(\tilde{R}_{i,N} - R_f)} = \\
&= \frac{cov(\tilde{R}_{i,N}, \tilde{R}_{P,N})}{var(\tilde{R}_{i,N})} = \\
&= \frac{cov((1 + \tilde{R}_{i,B})(1 - \phi_i) - 1, \tilde{R}_{P,N})}{var((1 + \tilde{R}_{i,B})(1 - \phi_i) - 1)} = \\
&= \frac{(1 - \phi_i)cov(\tilde{R}_{i,B}, \tilde{R}_{P,N})}{(1 - \phi_i)^2 var(\tilde{R}_{i,B})} = \\
&= \frac{cov(\tilde{R}_{i,B}, \tilde{R}_{P,N})}{(1 - \phi_i)var(\tilde{R}_{i,B})}.
\end{aligned}
$$

So for benchmark i, $\beta_{P,N,N,i} = \frac{\beta_{P,N,B,i}}{(1-\phi_i)}$ or $\beta_{P,N,B,i} = (1 - \phi_i)\beta_{P,N,N,i}$.

Placing that result in (*)

$$
\begin{aligned}
(*)\alpha_{P,N,N} - \alpha_{P,N,B} &= \sum_{i=1}^{N}[\beta_{P,N,B,i}(\bar{R}_{i,B} - R_f) - \beta_{P,N,N,i}(\bar{R}_{i,N} - R_f)] = \\
&= \sum_{i=1}^{N}[\beta_{P,N,B,i}(\bar{R}_{i,B} - R_f) - \frac{\beta_{P,N,B,i}}{(1 - \phi_i)}(\bar{R}_{i,N} - R_f)] = \\
&= \sum_{i=1}^{N}\beta_{P,N,B,i}(\bar{R}_{i,B} - R_f - \frac{\bar{R}_{i,N}}{1 - \phi_i} + \frac{R_f}{1 - \phi_i}).
\end{aligned}
$$

Recall that $\bar{R}_{i,N} = (1+\bar{R}_{i,B})(1-\phi_i) - 1$, then

$$\begin{aligned}
(*)\alpha_{P,N,N} - \alpha_{P,N,B} &= \sum_{i=1}^{N} \beta_{P,N,B,i}(\bar{R}_{i,B} - R_f - \frac{(1+\bar{R}_{i,B})(1-\phi_i)-1}{1-\phi_i} + \frac{R_f}{1-\phi_i}) = \\
&= \sum_{i=1}^{N} \beta_{P,N,B,i}(\bar{R}_{i,B} - R_f - 1 - \bar{R}_{i,B} + \frac{1}{1-\phi_i} + \frac{R_f}{1-\phi_i}) = \\
&= \sum_{i=1}^{N} \beta_{P,N,B,i}[(1+R_f)(\frac{1}{1-\phi_i} - 1)] = \\
&= \sum_{i=1}^{N} \beta_{P,N,B,i}[(1+R_f)(\frac{\phi_i}{1-\phi_i})] = \\
&= \sum_{i=1}^{N} [\beta_{P,N,B,i}(\frac{\phi_i}{1-\phi_i}) + R_f\beta_{P,N,B,i}(\frac{\phi_i}{1-\phi_i})].
\end{aligned}$$

Isolating $\alpha_{P,N,N}$ in term of $\alpha_{P,N,B}$ yields

$$\alpha_{P,N,N} = \alpha_{P,N,B} + \sum_{i=1}^{N} [\beta_{P,N,B,i}(\frac{\phi_i}{1-\phi_i}) + R_f\beta_{P,N,B,i}(\frac{\phi_i}{1-\phi_i})]. \tag{11}$$

Since $\beta_{P,N,B,i}(\frac{\phi_i}{1-\phi_i}) \gg R_f\beta_{P,N,B,i}(\frac{\phi_i}{1-\phi_i})$, then

$$\alpha_{P,N,N} \approx \alpha_{P,N,B} + \sum_{i=1}^{N} \beta_{P,N,B,i}(\frac{\phi_i}{1-\phi_i}) \approx \alpha_{P,N,B} + \sum_{i=1}^{N} \beta_{P,N,B,i}\phi_i = \alpha_{NET}. \tag{12}$$

We obtain approximately the same expression as in section 2 analysis. Hence, for positive betas, our net economic alpha is approximately and practically $\alpha_{N,N}$ - fund P's alpha as extracted from a regression of the net (after fees) fund return on the net (after fees) benchmarks returns.

Next, assume that all betas are negative. Thus, $\tilde{R}_{i,N} = (1+\tilde{R}_{i,B})(1+\phi_i) - 1$. For benchmark i, $\beta_{P,N,B,i} = \frac{cov(\tilde{R}_{i,B}-R_f, \tilde{R}_{P,N}-R_f)}{var(\tilde{R}_{i,B}-R_f)}$ and $\beta_{P,N,N,i} = \frac{cov(\tilde{R}_{i,N}-R_f, \tilde{R}_{P,N}-R_f)}{var(\tilde{R}_{i,N}-R_f)}$. Thus,

$$\begin{aligned}
\beta_{P,N,N,i} &= \frac{cov(\tilde{R}_{i,N} - R_f, \tilde{R}_{P,N} - R_f)}{var(\tilde{R}_{i,N} - R_f)} = \\
&= \frac{cov(\tilde{R}_{i,N}, \tilde{R}_{P,N})}{var(\tilde{R}_{i,N})} = \\
&= \frac{cov((1+\tilde{R}_{i,B})(1+\phi_i) - 1, \tilde{R}_{P,N})}{var((1+\tilde{R}_{i,B})(1+\phi_i) - 1)} = \\
&= \frac{(1+\phi_i)cov(\tilde{R}_{i,B}, \tilde{R}_{P,N})}{(1+\phi_i)^2 var(\tilde{R}_{i,B})} =
\end{aligned}$$

$$= \frac{cov(\tilde{R}_{i,B}, \tilde{R}_{P,N})}{(1+\phi_i)var(\tilde{R}_{i,B})}.$$

So for benchmark i, $\beta_{P,N,N,i} = \frac{\beta_{P,N,B,i}}{(1+\phi_i)}$ or $\beta_{P,N,B,i} = (1+\phi_i)\beta_{P,N,N,i}$.

Placing that result in (*)

$$\begin{aligned}
(*)\alpha_{P,N,N} - \alpha_{P,N,B} &= \sum_{i=1}^{N}[\beta_{P,N,B,i}(\bar{R}_{i,B} - R_f) - \beta_{P,N,N,i}(\bar{R}_{i,N} - R_f)] = \\
&= \sum_{i=1}^{N}[\beta_{P,N,B,i}(\bar{R}_{i,B} - R_f) - \frac{\beta_{P,N,B,i}}{(1+\phi_i)}(\bar{R}_{i,N} - R_f)] = \\
&= \sum_{i=1}^{N}\beta_{P,N,B,i}(\bar{R}_{i,B} - R_f - \frac{\bar{R}_{i,N}}{1+\phi_i} + \frac{R_f}{1+\phi_i}).
\end{aligned}$$

Recall that $\bar{R}_{i,N} = (1+\bar{R}_{i,B})(1+\phi_i) - 1$, then

$$\begin{aligned}
(*)\alpha_{P,N,N} - \alpha_{P,N,B} &= \sum_{i=1}^{N}\beta_{P,N,B,i}(\bar{R}_{i,B} - R_f - \frac{(1+\tilde{R}_{i,B})(1+\phi_i) - 1}{1+\phi_i} + \frac{R_f}{1+\phi_i}) = \\
&= \sum_{i=1}^{N}\beta_{P,N,B,i}(\bar{R}_{i,B} - R_f - 1 - \bar{R}_{i,B} + \frac{1}{1+\phi_i} + \frac{R_f}{1+\phi_i}) = \\
&= \sum_{i=1}^{N}\beta_{P,N,B,i}[(1+R_f)(\frac{1}{1+\phi_i} - 1)] = \\
&= \sum_{i=1}^{N}\beta_{P,N,B,i}[(1+R_f)(\frac{-\phi_i}{1+\phi_i})] = \\
&= \sum_{i=1}^{N}[\beta_{P,N,B,i}(\frac{-\phi_i}{1+\phi_i}) + R_f\beta_{P,N,B,i}(\frac{-\phi_i}{1+\phi_i})].
\end{aligned}$$

Isolating $\alpha_{P,N,N}$ in term of $\alpha_{P,N,B}$ yields

$$\alpha_{P,N,N} = \alpha_{P,N,B} + \sum_{i=1}^{N}[\beta_{P,N,B,i}(\frac{-\phi_i}{1+\phi_i}) + R_f\beta_{P,N,B,i}(\frac{-\phi_i}{1+\phi_i})]. \tag{13}$$

Since $|\beta_{P,N,B,i}(\frac{-\phi_i}{1+\phi_i})| \gg |R_f\beta_{P,N,B,i}(\frac{-\phi_i}{1+\phi_i})|$, then

$$\alpha_{P,N,N} \approx \alpha_{P,N,B} + \sum_{i=1}^{N}\beta_{P,N,B,i}(\frac{-\phi_i}{1+\phi_i}) \approx \alpha_{P,N,B} + \sum_{i=1}^{N}\beta_{P,N,B,i}(-\phi_i) = \alpha_{NET}. \tag{14}$$

We obtain approximately the same expression as in section 2 analysis. Hence, for negative betas, our net economic alpha is approximately and practically $\alpha_{N,N}$.

Beta greater than zero implies $\alpha_{NET} = \alpha_{P,N,B} + \sum_{i=1}^{N}\beta_{P,N,B,i}\phi_i$ and beta lower than zero implies $\alpha_{NET} = \alpha_{P,N,B} + \sum_{i=1}^{N}\beta_{P,N,B,i}(-\phi_i)$. Thus, we can write $\alpha_{NET} =$

$\alpha_{P,N,B}+\sum_{i=1}^{N}|\beta_{P,N,B,i}\phi_i| = \alpha_{P,N,B}+\sum_{i=1}^{N}|\beta_{P,N,B,i}|\phi_i$. α_{NET} is now approximately and practically $\alpha_{N,N}$ for both positive and negative betas.

APPENDIX 3: Estimating The Expense Ratio of ETFs

Table: A List of ETFs Following Various Investment Styles and Asset Pricing Factors

Name	Inception	Expense ratio (%)	Benchmark	Details	Long/Short	Link
iShares S&P 500 Index Fund	15/5/2000	0.09	Style	S&P500	long	http://us.ishares.com/product_info/fund/overview/IVV.htm
ProShares Short S&P500	19/6/2006	0.92	Style	S&P500	short	http://www.proshares.com/funds/sh.html
iShares Russell 1000 Index Fund	15/5/2000	0.15	Style	Russell 1000	long	http://us.ishares.com/product_info/fund/overview/IWB.htm
Direxion Daily Large Cap Bear 3x Shares	11/5/2008	0.95	Style	Russell 1000	short (-X3)	http://www.direxionshares.com/etf/lc_bear_3x_shares.html
iShares Russell 1000 Growth Index Fund	22/5/2000	0.2	Style	Russell 1000 Growth	long	http://us.ishares.com/product_info/fund/overview/IWF.htm
ProShares UltraShort Russell1000 Growth	20/2/2007	0.95	Style	Russell 1000 Growth	short (-X2)	http://www.proshares.com/funds/sfk.html
iShares Russell 1000 Value Index Fund	22/5/2000	0.2	Style	Russell 1000 Value	long	http://us.ishares.com/product_info/fund/overview/IWF.htm
ProShares UltraShort Russell1000 Value	20/2/2007	0.95	Style	Russell 1000 Value	short (-X2)	http://www.proshares.com/funds/sjf.html
iShares Russell Midcap Index Fund	17/7/2001	0.21	Style	Russell Midcap	long	http://us.ishares.com/product_info/fund/overview/IWR.htm
Direxion Daily Mid Cap Bear 3x Shares	1/8/2009	0.95	Style	Russell Midcap	short (-X3)	http://www.direxionshares.com/etf/mc_bear_3x_shares.html
iShares Russell Midcap Growth Index Fund	17/7/2001	0.25	Style	Russell Midcap Growth	long	http://us.ishares.com/product_info/fund/overview/IWP.htm
ProShares UltraShort Russell MidCap Growth	20/2/2007	0.95	Style	Russell MidCap Growth	short (-X2)	http://www.proshares.com/funds/sdk.html
iShares Russell Midcap Value Index Fund	17/7/2001	0.26	Style	Russell Midcap Value	long	http://us.ishares.com/product_info/fund/overview/IWS.htm
ProShares UltraShort Russell MidCap Value	20/2/2007	0.95	Style	Russell MidCap Value	short (-X2)	http://www.proshares.com/funds/sjl.html
iShares Russell 2000 Index Fund	22/5/2000	0.26	Style	Russell 2000	long	http://us.ishares.com/product_info/fund/overview/IWM.htm
ProShares Short Russell2000	23/1/2007	0.95	Style	Russell 2000	short	http://www.proshares.com/funds/rwm.html
iShares Russell 2000 Growth Index Fund	24/7/2000	0.25	Style	Russell 2000 Growth	long	http://us.ishares.com/product_info/fund/overview/IWO.htm
ProShares UltraShort Russell2000 Growth	20/2/2007	0.95	Style	Russell 2000 Growth	short (-X2)	http://www.proshares.com/funds/skk.html
iShares Russell 2000 Value Index Fund	24/7/2000	0.4	Style	Russell 2000 Value	long	http://us.ishares.com/product_info/fund/overview/IWN.htm
ProShares UltraShort Russell2000 Value	20/2/2007	0.95	Style	Russell 2000 Value	short (-X2)	http://www.proshares.com/funds/sjh.html
iShares Russell 2000 Index Fund	22/5/2000	0.26	SMB	Small Cap	long	http://us.ishares.com/product_info/fund/overview/IWM.htm
ProShares Short Russell2000	23/1/2007	0.95	SMB	Small Cap	short	http://www.proshares.com/funds/rwm.html
iShares Russell 1000 Index Fund	15/5/2000	0.15	SMB	Large Cap	long	http://us.ishares.com/product_info/fund/overview/IWB.htm
Direxion Daily Large Cap Bear 3x Shares	11/5/2008	0.95	SMB	Large Cap	short (-X3)	http://www.direxionshares.com/etf/lc_bear_3x_shares.html
iShares Russell 3000 Value Index Fund	24/7/2000	0.25	HML	Value	long	http://us.ishares.com/product_info/fund/overview/IWW.htm
ProShares UltraShort Russell1000 Value	20/2/2007	0.95	HML	Value	short (-X2)	http://www.proshares.com/funds/sjf.html
ProShares UltraShort Russell2000 Value	20/2/2007	0.95	HML	Value	short (-X2)	http://www.proshares.com/funds/sjh.html
iShres Russell 3000 Growth Index Fund	24/7/2000	0.25	HML	Growth	long	http://us.ishares.com/product_info/fund/overview/IWZ.htm
ProShares UltraShort Russell1000 Growth	20/2/2007	0.95	HML	Growth	short (-X2)	http://www.proshares.com/funds/sfk.html
ProShares UltraShort Russell2000 Growth	20/2/2007	0.95	HML	Growth	short (-X2)	http://www.proshares.com/funds/skk.html
PowerShares DWA Technical Leaders Portfolio	1/3/2007	0.6	Momentum		long	http://www.invescopowershares.com/products/overview.aspx?ticker=PDP
JETS Contrarian Opportunities Index Fund	7/4/2010	0.58	Momentum		short	http://www.javelinfunds.com/jco-djim.php

Table 1: Data Summary

Table 1 reports the summary characteristics of the data used in this paper. The data comprise nine years (2001-2009) of non-specialized open-end equity funds for twelve Lipper classifications: LCCE (Large-Cap Core Funds), MCCE (Mid-Cap Core Funds), MLCE (Multi-Cap Core Funds), SCCE (Small-Cap Core Funds), LCGE (Large-Cap Growth Funds), MCGE (Mid-Cap Growth Funds), MLGE (Multi-Cap Growth Funds),SCGE (Small-Cap Growth Funds), LCVE (Large-Cap Value Funds), MCVE (Mid-Cap Value Funds), MLVE (Multi-Cap Value Funds), and SCVE (Small-Cap Value Funds).

Panels A, B, and C specify, per each classification, the average number of observations, the average of the Total-Net-Assets (TNA - in millions of dollars), and the average and weighted (by TNA) average of the expense ratios and management fees charged by funds managers. Panel A reports the 2001-2003 data summary, Panel B reports the 2004-2006 data summary, and Panel C reports the 2007-2009 data summary. Panel D reports, per each classification, the average distribution of the style benchmarks (the style benchmarks are Russell 1000, Russell 1000 growth, Russell 1000 value, Russell midcap, Russell midcap growth, Russell midcap value, Russell 2000, Russell 2000 growth, Russell 2000 value, and S&P500).

Panel A: Data Characteristics 2001-2003

			AVERAGE		WEIGHTED	
CLASS	OBS	TNA	EXP RATIO	MGMT FEE	EXP RATIO	MGMT FEE
LCCE	188	1,168	1.20%	0.64%	1.12%	0.65%
LCGE	123	843	1.38%	0.71%	1.29%	0.72%
LCVE	66	1,259	1.18%	0.63%	1.06%	0.63%
MCCE	52	544	1.10%	0.63%	1.00%	0.63%
MCGE	85	311	1.46%	0.81%	1.37%	0.81%
MCVE	34	344	1.38%	0.83%	1.32%	0.82%
MLCE	68	878	1.07%	0.59%	1.09%	0.59%
MLGE	68	947	1.58%	0.84%	1.51%	0.84%
MLVE	98	886	1.24%	0.68%	1.17%	0.68%
SCCE	89	286	1.23%	0.71%	1.15%	0.71%
SCGE	100	217	1.51%	0.88%	1.45%	0.89%
SCVE	66	332	1.35%	0.88%	1.30%	0.88%
AVG:		667.9	1.31%	0.74%	1.24%	0.74%

Table 1 - Cont.

Panel B: Data Carachteristics 2004-2006

		AVERAGE			WEIGHTED	
CLASS	OBS	TNA	EXP RATIO	MGMT FEE	EXP RATIO	MGMT FEE
LCCE	223	1,245	1.20%	0.64%	1.09%	0.64%
LCGE	146	813	1.30%	0.70%	1.18%	0.70%
LCVE	90	1,598	1.22%	0.65%	1.09%	0.65%
MCCE	95	915	1.94%	0.98%	1.85%	0.98%
MCGE	129	475	1.46%	0.79%	1.34%	0.79%
MCVE	67	1,333	1.28%	0.79%	1.23%	0.79%
MLCE	192	789	1.23%	0.68%	1.19%	0.68%
MLGE	99	1,969	1.44%	0.77%	1.36%	0.77%
MLVE	126	1,165	1.16%	0.67%	1.08%	0.66%
SCCE	177	550	1.35%	0.83%	1.27%	0.84%
SCGE	139	357	1.54%	0.93%	1.44%	0.93%
SCVE	64	445	1.39%	0.89%	1.28%	0.90%
AVG:		971.2	1.37%	0.78%	1.28%	0.78%

Panel C: Data Characteristics 2007-2009

		AVERAGE			WEIGHTED	
CLASS	OBS	TNA	EXP RATIO	MGMT FEE	EXP RATIO	MGMT FEE
LCCE	168	822	1.04%	0.60%	0.95%	0.60%
LCGE	145	1,102	1.22%	0.71%	1.07%	0.71%
LCVE	91	1,392	1.13%	0.65%	1.00%	0.64%
MCCE	72	677	1.18%	0.74%	1.07%	0.74%
MCGE	105	411	1.33%	0.76%	1.21%	0.76%
MCVE	48	601	1.27%	0.77%	1.17%	0.76%
MLCE	163	934	1.08%	0.65%	1.04%	0.65%
MLGE	86	1,427	1.28%	0.74%	1.19%	0.74%
MLVE	80	611	1.12%	0.67%	1.03%	0.67%
SCCE	166	486	1.27%	0.78%	1.15%	0.78%
SCGE	128	323	1.52%	0.88%	1.40%	0.88%
SCVE	60	332	1.35%	0.84%	1.22%	0.84%
AVG:		759.6	1.23%	0.73%	1.12%	0.73%

Table 1 - Cont.

Panel D: The Relation Between LIPPER Classification and Fund's Style (most correlated benchmark)

The Table reports, for each Lipper classification, the percentage of funds in each classification that are most correlated with Russell 1000, Russell 1000 growth, Russell 1000 value, Russell midcap, Russell midcap growth, Russell midcap value, Russell 2000, Russell 2000 growth, Russell 2000 value, and S&P500. Each raw sums up to 100% approximately.

	PERCENTAGE OF FUNDS MOST CORRELATED WITH									
LIPPER	RSL 1000			RSL MIDCAP			RSL 2000			S&P500
CLASSIFICATION	GENERAL	GROWTH	VALUE	GENERAL	GROWTH	VALUE	GENERAL	GROWTH	VALUE	
LCCE	38.1%	9.7%	3.2%	1.4%	1.8%	1.1%	0.0%	0.3%	0.0%	44.2%
LCGE	8.6%	72.4%	0.2%	3.6%	11.8%	0.0%	0.0%	0.2%	0.0%	3.1%
LCVE	9.7%	0.4%	70.4%	1.2%	0.0%	0.4%	0.0%	0.0%	0.4%	17.6%
MCCE	1.4%	0.5%	0.0%	64.9%	11.0%	5.9%	5.8%	8.8%	1.7%	0.0%
MCGE	0.3%	1.9%	0.3%	10.6%	73.9%	0.8%	1.7%	10.3%	0.3%	0.0%
MCVE	2.7%	0.7%	0.7%	42.4%	3.4%	38.9%	4.9%	2.9%	3.4%	0.0%
MLCE	44.4%	8.0%	4.8%	22.0%	6.5%	3.5%	1.5%	1.7%	0.6%	7.1%
MLGE	7.2%	31.7%	0.0%	7.9%	46.0%	1.4%	1.2%	3.5%	0.0%	1.1%
MLVE	18.7%	0.4%	40.3%	11.7%	1.7%	14.4%	1.6%	0.9%	0.4%	9.9%
SCCE	0.0%	0.2%	0.0%	5.3%	1.6%	2.6%	64.3%	17.5%	8.5%	0.0%
SCGE	0.0%	0.8%	0.0%	3.6%	7.8%	0.0%	9.2%	78.2%	0.5%	0.0%
SCVE	0.0%	0.0%	1.1%	2.2%	1.1%	9.3%	30.7%	2.2%	53.4%	0.0%
AVG:	10.9%	10.6%	10.1%	14.7%	13.9%	6.5%	10.1%	10.5%	5.8%	6.9%

Table 2: Nine Benchmark-Models of Funds Return

The benchmark models are: 1. M: the CAPM - the market factor, 2. S: the fund's style - the fund's style is determined by identifying the benchmark that generates the highest correlation with respect to the fund's return, 3. G(Class): the group factor - the average return of funds belonging to the same group. The funds are gathered into groups based on funds' Lipper classification, 4. G(Style): the group factor - the average return of funds belonging to the same group. The funds are gathered into groups based on funds' style, 5. 4F: Carhart's four-factor model, 6. 4F-S: a five-factor model, based on Hunter et al methodology, that includes Carhart's four-factor model and the style factor, 7. 4F-G(Class): Hunter et al five-factor model (Carhart's four-factor model plus the group factor) - funds are gathered into groups based on funds' Lipper classification, 8. 4F-G(Style): Hunter et al five-factor model (Carhart's four-factor model plus the group factor) - funds are gathered into groups based on funds' style, and 9. 3F: Fama and French three-factor model. We use all funds that exist throughout all 36 months of each period. We denote $\bar{\beta}_{Style}$ as the average factor loading of the style factor, $\bar{\beta}_{Group}$ as the average factor loading of the group factor, $\bar{\beta}_{Market}$ as the average factor loading of the market factor, $\bar{\beta}_{SMB}$ as the average factor loading of the SMB (Small Minus Big) factor, $\bar{\beta}_{HML}$ as the average factor loading of the HML (High Minus Low) factor, $\bar{\beta}_{MOM}$ as the average factor loading of the momentum factor, $\bar{\alpha}$ as the average standard alpha, and $\bar{\alpha}_{NET}$ as the average net economic alpha.

$\bar{\alpha}$ is computed as follows: first, for each fund, the intercept from the regression of net of fees fund excess return on the excess return on the appropriate indexes is extracted. Second, the fund's annualized alpha is calculated as $(1+intercept)^{12}-1$. Then we take the average of the annualized alpha of all funds and obtain $\bar{\alpha}$. $\bar{\alpha}_{NET}$, is computed as follows: for each fund P, we add to the annualized alpha, α_P, the cost of mimicking the fund's systematic risk via ETFs: $\alpha_{NET,P} = \alpha_P + \sum_{i=1}^{N} \phi_i |\beta_{P,i}|$, where ϕ_i is the expense ratio for benchmark i and $\beta_{P,i}, i = 1..N$ are extracted from the regression of net of fees fund excess return on the excess return on the appropriate indexes. We then take the average of the net economic alpha of all funds and obtain $\bar{\alpha}_{NET}$.

For each of the nine benchmark models, Panel A reports the 2001-2003 regression results, Panel B reports the 2004-2006 regression results, Panel C reports the 2007-2009 regression results, and Panel D reports the average regression results of the three periods (2001-2003, 2004-2006 and 2007-2009).

Table 2 - Cont.

Panel A: 2001-2003 Regressions Summary

MODEL	OBS	R-squared	$\bar{\beta}_{Style}$	$\bar{\beta}_{Group}$	$\bar{\beta}_{Market}$	$\bar{\beta}_{SMB}$	$\bar{\beta}_{HML}$	$\bar{\beta}_{MOM}$	ANNUAL $\bar{\alpha}$	ANNUAL $\bar{\alpha}_{NET}$
M	1,037	0.83			1.00				0.88%**	0.97%**
S	1,037	0.92	0.95						-1.36%**	-1.16%**
G(Class)	1,037	0.93		1.00					N.R	N.R
G(Style)	1,037	0.93		1.00					N.R	N.R
3F	1,037	0.91			1.00	0.24	0.09		-3.79%**	-2.99%**
4F	1,037	0.93			1.05	0.26	0.05	0.06	-3.68%**	-2.66%**
4F-S	1,037	0.94	0.74		1.05	0.26	0.05	0.06	-2.25%**	-1.02%**
4F-G(Class)	1,037	0.95		1.00	1.05	0.26	0.05	0.06	N.R	N.R
4F-G(Style)	1,037	0.95		1.00	1.05	0.26	0.05	0.06	N.R	N.R

N.R - not relevant. In models including the group factor (G), mean alpha is zero by definition (see Appendix 2)

**The coefficient is different from zero at the 5% significance level.

**The coefficient is different from zero at the 1% significance level.

We also test the null hypothesis H0: $\alpha = \alpha_{NET}$, and find that alpha is different from the net economic alpha at the 1% significance level for all testable benchmark-models.

Panel B: 2004-2006 Regressions Summary

MODEL	OBS	R-squared	$\bar{\beta}_{Style}$	$\bar{\beta}_{Group}$	$\bar{\beta}_{Market}$	$\bar{\beta}_{SMB}$	$\bar{\beta}_{HML}$	$\bar{\beta}_{MOM}$	ANNUAL $\bar{\alpha}$	ANNUAL $\bar{\alpha}_{NET}$
M	1,547	0.81			1.19				-2.50%**	-2.39%**
S	1,547	0.87	0.96						-0.98%**	-0.77%**
G(Class)	1,547	0.89		1.00					N.R	N.R
G(Style)	1,547	0.88		1.00					N.R	N.R
3F	1,547	0.88			0.96	0.28	-0.03		-0.70%**	0.12%
4F	1,547	0.89			0.96	0.23	-0.09	0.09	-0.30%**	0.66%**
4F-S	1,547	0.90	0.66		0.96	0.23	-0.09	0.09	-0.81%**	0.33%*
4F-G(Class)	1,547	0.91		1.00	0.96	0.23	-0.09	0.09	N.R	N.R
4F-G(Style)	1,547	0.91		1.00	0.96	0.23	-0.09	0.09	N.R	N.R

N.R - not relevant. In models including the group factor (G), mean alpha is zero by definition (see Appendix 2)

**The coefficient is different from zero at the 5% significance level.

**The coefficient is different from zero at the 1% significance level.

We also test the null hypothesis H0: $\alpha = \alpha_{NET}$, and find that alpha is different from the net economic alpha at the 1% significance level for all testable benchmark-models.

Table 2 - Cont.

Panel C: 2007-2009 Regressions Summary

MODEL	OBS	R-squared	$\bar{\beta}_{Style}$	$\bar{\beta}_{Group}$	$\bar{\beta}_{Market}$	$\bar{\beta}_{SMB}$	$\bar{\beta}_{HML}$	$\bar{\beta}_{MOM}$	ANNUAL $\bar{\alpha}$	ANNUAL $\bar{\alpha}_{NET}$
M	1,312	0.90			1.01				0.23%*	0.32%**
S	1,312	0.94	0.97						-0.13%	0.07%
G(Class)	1,312	0.95		1.00					N.R	N.R
G(Style)	1,312	0.94		1.00					N.R	N.R
3F	1,312	0.94			1.00	0.26	-0.08		-0.71%**	0.08%
4F	1,312	0.95			1.00	0.26	-0.08	(0.00)	-0.71%**	0.13%
4F-S	1,312	0.96	0.71		1.00	0.26	-0.08	(0.00)	-0.36%**	0.66%**
4F-G(Class)	1,312	0.96		1.00	1.00	0.26	-0.08	(0.00)	N.R	N.R
4F-G(Style)	1,312	0.96		1.00	1.00	0.26	-0.08	(0.00)	N.R	N.R

N.R - not relevant. In models including the group factor (G), mean alpha is zero by definition (see Appendix 2)

**The coefficient is different from zero at the 5% significance level.

**The coefficient is different from zero at the 1% significance level.

We also test the null hypothesis H0: $\alpha = \alpha_{NET}$, and find that alpha is different from the net economic alpha at the 1% significance level for all testable benchmark-models.

Panel D: 2001-2003, 2004-2006 and 2007-2009 Average Regressions Summary

MODEL	OBS	R-squared	$\bar{\beta}_{Style}$	$\bar{\beta}_{Group}$	$\bar{\beta}_{Market}$	$\bar{\beta}_{SMB}$	$\bar{\beta}_{HML}$	$\bar{\beta}_{MOM}$	ANNUAL $\bar{\alpha}$	ANNUAL $\bar{\alpha}_{NET}$
M	1,299	0.85			1.07				-0.46%	-0.37%
S	1,299	0.91	0.96						-0.82%	-0.62%
G(Class)	1,299	0.92		1.00					N.R	N.R
G(Style)	1,299	0.91		1.00					N.R	N.R
3F	1,299	0.91			0.98	0.26	-0.01		-1.73%	-0.93%
4F	1,299	0.92			1.00	0.25	-0.04	0.05	-1.56%	-0.62%
4F-S	1,299	0.93	0.70		1.00	0.25	-0.04	0.05	-1.14%	-0.01%
4F-G(Class)	1,299	0.94		1.00	1.00	0.25	-0.04	0.05	N.R	N.R
4F-G(Style)	1,299	0.94		1.00	1.00	0.25	-0.04	0.05	N.R	N.R

Table 3: Correlations Between Nine Benchmark Models Alphas

Table 3 reports the average correlations between funds alphas. First, alphas are estimated for each benchmark model in each of the subperiods (2001-2003, 2004-2006, 2007-2009), using all funds that exist throughout all 36 months of each period. Then, the correlations between the different benchmark model alphas are calculated.

	CORRELATIONS BETWEEN THE FOLLOWING BENCHMARK MODELS' ALPHAS								
BENCHMARK MODEL	M	S	G(Class)	G(Style)	3F	4F	4F-S	4F-G(Class)	4F-G(Style)
M	1	0.74	0.63	0.68	0.56	0.53	0.61	0.48	0.56
S	0.74	1	0.78	0.92	0.81	0.77	0.83	0.71	0.76
G(Class)	0.63	0.78	1	0.83	0.73	0.71	0.74	0.78	0.76
G(Style)	0.68	0.92	0.83	1	0.79	0.76	0.77	0.77	0.83
3F	0.56	0.81	0.73	0.79	1	0.98	0.89	0.78	0.80
4F	0.53	0.77	0.71	0.76	0.98	1	0.89	0.79	0.80
4F-S	0.61	0.83	0.74	0.77	0.89	0.89	1	0.85	0.87
4F-G(Class)	0.48	0.71	0.78	0.77	0.78	0.79	0.85	1	0.92
4F-G(Style)	0.56	0.76	0.76	0.83	0.80	0.80	0.87	0.92	1

See Table 2 for benchmark model definition.

Table 4: Differences in Classification and Ranking Across Nine Possible Benchmark-Models

Table 4 compares the results implied by the nine benchmark models.

The nine benchmark models (specified at Table 2) are applied for 2001-2003, 2004-2007 and 2007-2009. We use all funds that exist throughout all 36 months of each period.

Based on the different alphas of the different benchmark models, we classify and rank the funds. We then compare funds classification and ranking and report the 2001-2003, 2004-2006 and 2007-2009 average comparison results.

Panel A: Classification Differences Among The Benchmark Models

Based on each alpha, a fund is classified as either a good fund (alpha is greater than zero) or a bad fund (alpha is lower than zero). If a fund is classified as a good (bad) fund according to both benchmark model i and benchmark model j, then there is no classification difference between models i and j. On the other hand, if a fund is classified as a good (bad) fund according to benchmark model i and is classified as a bad (good) fund according to benchmark model j, than there is a classification difference between models i and j. Panel A reports the 2001-2003, 2004-2006 and 2007-2009 average frequency of classification differences between the benchmark models.

	M	S	G(Class)	G(Style)	4F	4F-S	4F-G(CLASS)	4F-G(STYLE)	3F
M	0.0%	24.9%	31.3%	27.7%	30.8%	28.1%	34.1%	30.2%	28.8%
S	24.9%	0.0%	25.7%	17.0%	25.0%	18.9%	26.5%	23.5%	21.9%
G(Class)	31.3%	25.7%	0.0%	17.5%	27.9%	27.7%	21.2%	23.1%	27.1%
G(Style)	27.7%	17.0%	17.5%	0.0%	27.0%	25.7%	23.0%	20.0%	24.7%
3F	28.8%	21.9%	27.1%	24.7%	6.0%	13.0%	22.4%	22.1%	0.0%
4F	30.8%	25.0%	27.9%	27.0%	0.0%	12.5%	21.9%	21.4%	6.0%
4F-S	28.1%	18.9%	27.7%	25.7%	12.5%	0.0%	19.5%	17.0%	13.0%
4F-G(Class)	34.1%	26.5%	21.2%	23.0%	21.9%	19.5%	0.0%	13.0%	22.4%
4F-G(Style)	30.2%	23.5%	23.1%	20.0%	21.4%	17.0%	13.0%	0.0%	22.1%
AVG*:	29.5%	22.9%	25.2%	22.8%	21.6%	20.3%	22.7%	21.3%	20.7%

Under the null that classification is random and independent across benchmark models difference proportion should equal 0.5. We test the null hypothesis that the differences proportions equal 0.5 and find that all difference proportions in the Table are different from 0.5 at the 1% significance level.

*AVG do not include signs differences between model i and itself

Table 4 - Cont.

Panel B: Rank Differences (of more than a decile) Among The Benchmark Models

For each benchmark model, the best performing fund with the highest alpha is ranked at the 1st place, the second best performing fund with the second highest alpha is ranked at the 2nd place, and so on. If fund P is ranked in the k-th place according to benchmark model i and is ranked in the l-th place according to benchmark model j, then if $|k-l| < 0.1N$ then there is no ranking difference between models i and j. On the other hand, If $|k-l| \geq 0.1N$, then there is a ranking difference between models i and j. We compute ranking differences in each subperiod (2001-2003, 2004-2006 and 2007-2009) and report averages across the 3 subperiods in Panel B.

	M	S	G(Class)	G(Style)	4F	4F-S	4F-G(CLASS)	4F-G(STYLE)	3F
M	0.0%	62.9%	57.1%	58.3%	59.8%	59.5%	64.5%	60.2%	56.5%
S	62.9%	0.0%	49.3%	29.2%	55.1%	44.4%	55.2%	50.8%	51.0%
G(Class)	57.1%	49.3%	0.0%	37.9%	56.8%	54.8%	49.3%	53.7%	54.6%
G(Style)	58.3%	29.2%	37.9%	0.0%	54.4%	49.5%	52.6%	45.0%	49.7%
3F	56.5%	51.0%	54.6%	49.7%	12.9%	36.4%	48.4%	45.5%	0.0%
4F	59.8%	55.1%	56.8%	54.4%	0.0%	31.8%	44.9%	45.0%	12.9%
4F-S	59.5%	44.4%	54.8%	49.5%	31.8%	0.0%	38.4%	33.4%	36.4%
4F-G(Class)	64.5%	55.2%	49.3%	52.6%	44.9%	38.4%	0.0%	29.6%	48.4%
4F-G(Style)	60.2%	50.8%	53.7%	45.0%	45.0%	33.4%	29.6%	0.0%	45.5%
AVG*:	59.8%	49.8%	51.7%	47.1%	45.1%	43.5%	47.9%	45.4%	44.4%

*AVG does not include differences between model i and itself

Panel C: Comparing Best 20 Performing Funds Among The Models

We narrow the data and keep only funds ranked as 20 best performing funds. Then, in each subperiod (2001-2003, 2004-2006 and 2007-2009), we count how many funds appear on both benchmark model i and benchmark model j lists of best performing funds. Averages of the 3 subperiods are reported in Panel C.

	M	S	G(Class)	G(Style)	4F	4F-S	4F-G(CLASS)	4F-G(STYLE)	3F
M	20.0	7.7	9.3	9.3	7.7	5.7	5.0	6.0	9.3
S	7.7	20.0	8.7	13.3	8.3	8.0	7.0	8.3	9.7
G(Class)	9.3	8.7	20.0	10.7	8.0	8.3	9.0	8.7	9.3
G(Style)	9.3	13.3	10.7	20.0	9.7	8.0	7.0	9.3	10.7
3F	9.3	9.7	9.3	10.7	16.0	11.7	10.7	10.0	20.0
4F	7.7	8.3	8.0	9.7	20.0	12.7	10.3	10.3	16.0
4F-S	5.7	8.0	8.3	8.0	12.7	20.0	12.7	13.3	11.7
4F-G(Class)	5.0	7.0	9.0	7.0	10.3	12.7	20.0	13.3	10.7
4F-G(Style)	6.0	8.3	8.7	9.3	10.3	13.3	13.3	20.0	10.0
AVG*:	7.5	8.9	9.0	9.8	10.4	10.0	9.4	9.9	10.9

*AVG does not include numbers on the diagonal (all of which are 20)

Table 4 - Cont.

Panel D: Comparing Worst 20 Performing Funds Among The Models

We keep only funds ranked as 20 worst performing funds. Then, in each subperiod (2001-2003, 2004-2006 and 2007-2009), we count how many funds appear on both benchmark model i and benchmark model j lists of worst performing funds. Averages of the 3 subperiods are reported in Panel D.

	M	S	G(Class)	G(Style)	4F	4F-S	4F-G(CLASS)	4F-G(STYLE)	3F
M	20.0	13.0	11.0	11.0	11.7	11.3	9.7	9.7	11.3
S	13.0	20.0	12.3	13.3	9.0	11.0	8.7	9.0	10.3
G(Class)	11.0	12.3	20.0	12.3	11.7	10.7	11.3	10.0	12.3
G(Style)	11.0	13.3	12.3	20.0	10.3	11.3	10.3	12.7	11.3
3F	11.3	10.3	12.3	11.3	17.3	13.7	12.3	13.7	20.0
4F	11.7	9.0	11.7	10.3	20.0	15.0	13.0	14.7	17.3
4F-S	11.3	11.0	10.7	11.3	15.0	20.0	14.0	15.3	13.7
4F-G(Class)	9.7	8.7	11.3	10.3	13.0	14.0	20.0	14.0	12.3
4F-G(Style)	9.7	9.0	10.0	12.7	14.7	15.3	14.0	20.0	13.7
AVG*:	11.1	10.8	11.5	11.6	12.8	12.8	11.7	12.4	12.8

*AVG does not include numbers on the diagonal (all of which are 20)

Table 5: The Impact of Net Economic Alphas on Classifications and Rankings

Table 5 reports, for each benchmark model (details about the benchmark models are provided in Table 2), the scoring differences between alphas and their corresponding net economic alphas. Differences are observed over 3 subperiods: 2001-2003, 2004-2007 and 2007-2009, and then an average is computed.

Panel A: Classification Differences - Alpha Vs. Net Economic Alpha

Panel A reports the frequency of differences in funds' classification between the alphas and the corresponding net economic alphas, i.e the frequency of cases where α is negative while α_{NET} is positive.

BENCHMARK MODEL	CLASSIFICATION DIFFERENCE FREQUENCY: α VS. α_{NET}
M	0.8%
S	3.0%
3F	9.5%
4F	8.7%
4F-S	12.7%
AVG:	6.9%

Panel B: Ranks Differences (of more than a decile) - Alpha Vs. Net Economic Alpha

Panel B reports the frequency of 10% differences in funds' rankings, when comparing alphas and their corresponding net economic alphas.

BENCHMARK MODEL	CONSIDERABLE RANKING DIFFERENCE FREQUENCY: α VS. α_{NET}
M	0.0%
S	0.0%
3F	2.3%
4F	2.4%
4F-S	4.0%
AVG:	1.7%

Table 5 - Cont.

Panel C: Comparing Best 20 Performing Funds - Alpha Vs. Net Economic Alpha

BENCHMARK MODEL	AVERAGE NUMBER OF IDENTICAL FUNDS IN THE BEST 20 PERFORMING FUNDS: α VS. α_{NET}
M	20.0
S	20.0
3F	17.7
4F	18.3
4F-S	17.7
AVG:	18.7

Panel D: Comparing Worst 20 Performing Funds - Alpha Vs. Net Economic Alpha

BENCHMARK MODEL	AVERAGE NUMBER OF IDENTICAL FUNDS IN THE WORST 20 PERFORMING FUNDS: α VS. α_{NET}
M	20.0
S	20.0
3F	18.3
4F	18.3
4F-S	17.3
AVG:	18.8

Table 6: Persistence of Funds Performance - Correlations of Subperiod Performance

Table 6 tests for persistence in funds performance for each of the nine benchmark models (the benchmark models are described in Table 2). We use all available funds and check for the correlation of various performance measures between the evaluation period (period 1) and the performance period (period 2). The following performance measures are examined: 1. funds' alphas, 2. funds' ranks (based on funds' alphas), where the ranks of the funds at the performance period include all funds existing at that period, whether they had existed at the evaluation period or not (original ranks), 3. funds' ranks (based on funds' alphas), where the ranks of the funds at the performance period include only surviving funds from the evaluation period (veteran surviving ranks), and 4. funds' deciles belongings (based on funds' alphas). Panel A reports the correlations between the 2001-2003 evaluation period and the 2004- 2006 performance period. Panel B reports the correlations between the 2004-2006 evaluation period and the 2007-2009 performance period. The correlations are calculated and reported for each of the nine benchmark models.

Panel A: Correlations Between The 2001-2003 Evaluation Period and The 2004-2006 Performance Period

	CORRELATION OF 2001-3 AND 2004-6				
BENCHMARK MODEL	ALPHA	ORIGINAL RANKS	SURVIVORS RANKS	DECILES RANKS	RAW AVG
M	0.28**	0.35**	0.35**	0.34**	0.33
S	0.24**	0.26**	0.26**	0.26**	0.26
G(Class)	0.19**	0.17**	0.17**	0.18**	0.18
G(Style)	0.11**	0.15**	0.14**	0.14**	0.14
3F	0.07	0.05	0.05	0.06	0.06
4F	-0.05	-0.06	-0.06	-0.06	-0.06
4F-S	-0.01	0.03	0.03	0.03	0.02
4F-G(Class)	-0.02	0.01	0.01	0.01	0.00
4F-G(Style)	0.02	0.08	0.08	0.08*	0.06
AVG	0.09	0.12	0.11	0.12	

*The coefficient is different from zero at the 5% significance level.

**The coefficient is different from zero at the 1% significance level.

Table 6 - Cont.

Panel B: Correlations Between The 2004-2006 Evaluation Period and The 2007-2009 Performance Period

	CORRELATION OF 2004-6 AND 2007-9				
BENCHMARK MODEL	ALPHA	ORIGINAL RANKS	SURVIVORS RANKS	DECILES RANKS	RAW AVG
M	-0.18**	-0.23**	-0.23**	-0.23**	-0.22
S	0.12**	0.14**	0.14**	0.14**	0.13
G(Class)	0.19**	0.18**	0.18**	0.18**	0.18
G(Style)	0.11**	0.11**	0.11**	0.11**	0.11
3F	0.25**	0.23**	0.23**	0.21**	0.23
4F	0.23**	0.20**	0.20**	0.20**	0.21
4F-S	0.18**	0.14**	0.14**	0.13**	0.15
4F-G(Class)	0.15**	0.14**	0.14**	0.14**	0.14
4F-G(Style)	0.18**	0.15**	0.15**	0.15**	0.16
AVG	0.14	0.12	0.12	0.11	

**The coefficient is different from zero at the 1% significance level.