2007

On Performance & Tracking Error in Exchange-Traded Funds and Index Mutual Funds

Steven Welch
College of Saint Benedict/Saint John's University, swelch@csbsju.edu

On Performance & Tracking Error in Exchange-Traded Funds and Index Mutual Funds

ABSTRACT

Using daily data, we find abnormal returns associated with the ETFs are higher than the alphas of the index funds in most cases. This result is in contrast to previous results that conclude that index funds tend to have higher alphas than ETFs. The results are much more prevalent in funds that follow the S&P 500 than funds that do not. One explanation for the difference in results is the more comprehensive sample of ETFs analyzed here. When looking at the components of abnormal returns, several regressions were performed. We find that market concentration, turnover, and no load are at least marginally significant for index funds and the constant, age, expense ratio, standard deviation (risk), and market concentration are significant determinants in the abnormal returns of ETFs. When examining the tracking errors for ETFs and index funds, we find index funds are able to track their indexes much better than ETFs and domestic ETFs are better than ETFs that track indexes in other countries. The most significant finding of this paper is that tracking error affects fund flow in the following period. While fund flows are generally increasing for both ETFs and index funds, funds that track their respective index better increase their net assets by a larger percentage than funds that track their index less well.

JEL Codes: G11, G14

Keywords: ETFs, index mutual funds, tracking error

This version:
September 1, 2007

I. INTRODUCTION

Exchange-traded funds (ETFs) and index mutual funds have similar objectives in that they attempt to closely match some index using a passive management strategy. However, the way that they are formed is somewhat different. Each one is organized in such a way that they have distinct cost structures and a slightly differing ability to accurately follow their given indexes.

Because ETFs and index mutual funds work differently but have similar objectives, it is interesting to examine the ability of each to match the performance of its index. Past studies suggest that while ETFs generally underperform their index fund counterparts in the short-term due to transactions costs, they overcome this shortcoming in the long-term by having slightly lower management expenses. However, ETFs have been very popular among short-term traders due to their trading advantages over mutual funds. Since ETFs tend to have relatively short histories, almost all studies look at a very limited data set – one or two ETFs and one or two index funds. This study expands on this limitation to look at several types of funds. By using funds that have not been studied in the past, we provide a more comprehensive idea of how ETFs and index mutual funds compare as a group rather than only looking at funds that track the S&P 500 or another popular index.

Using daily data, we find abnormal returns (alphas) associated with the ETFs are higher than the alphas of the index funds in most cases. This result directly contradicts results from previous literature. One explanation for this is that we look at a much larger sample of ETFs than do prior studies. Higher ETF alphas are much more prevalent in funds that follow the S&P 500 than

funds that do not. This finding may be the result of ETF managers improving their performance by timing their modifications to the fund in order to minimize transactions costs.[1]

We also look at which components of abnormal return are important determinants. We find confirming results that ETFs have higher abnormal returns than index funds. Other important factors for index funds include the relative size of a fund when compared with other funds following the same index and the Herfindahl index (given funds following a single index constitute an industry). For ETFs, the inception date, expense ratio, standard deviation, and relative size are all important determinants of abnormal return.

Another major issue when examining ETFs and index funds is "tracking error." The difference between an ETF or index fund's return and the return of the index that it is following is defined as its tracking error. Tracking error has several possible causes discussed in the next section. We examine the tracking error for both ETFs and index funds and compare them to each other to see which type of fund has less tracking error. Then, we determine if tracking error affects fund flow in subsequent periods. Our findings show that index funds are able to track their indexes much better than ETFs and domestic ETFs are better than ETFs that track indexes in other countries. The most significant finding of this paper is that tracking error affects fund flow in the following period. While fund flows are generally increasing for both ETFs and index funds, funds that track their respective index better increase their net assets by a larger percentage than funds that track their index less well. This result is important since it suggests that investors react to tracking errors.

[1] Gastineau (2004) says this method is common among index fund managers.

The rest of this paper is organized as follows: Section III delves into the previous literature on the topics analyzed here. The third section looks at the specific analyses and models used in this study. Section IV describes the data. The fifth section presents the empirical evidence, and the final section draws some conclusions.

II. LITERATURE REVIEW

Surprisingly, literature in the area of exchange-traded funds has not caught the attention yet of many academic researchers, and therefore, ETFs are not represented strongly in many top scholarly journals. There are numerous scholarly articles about mutual funds and index funds, however.

II.A. ETFs and Index Funds

Dellva (2001) performs a simulation exercise with one mutual fund and both ETFs that follow the S&P 500. He determined that small investors and short-term investors benefited most by investing in the mutual fund that he chose for his study. While helpful to investors that want to follow a broad market index, this study was very limited in its scope. Since only one index was used and three investment vehicles, drawing wide-ranging conclusions based upon study would be somewhat tenuous. Elton, Gruber, Comer and Li (2002) also studied Spiders and S&P 500 index funds. They find that Spiders underperform the index funds by about 18 basis points per year. They believe that the ability to trade Spiders throughout the day provides a value that more than outweighs this small discrepancy in performance. They explain that Spiders are often used as a risk control mechanism and for short-term trading. Also, even though they may offer less

return than some futures contracts, they have the advantage of being able to trade in much smaller increments. Poterba and Shoven (2002) examine the differences in returns between the SPDR trust and the Vanguard Index 500 fund. Both of these securities follow the S&P 500 index and have the advantage of being the largest ETF and the largest mutual fund. They show that ETFs perform virtually as well as index funds.

Kostovetsky (2003) shows that under any reasonable circumstances, a small investor would prefer an index mutual fund over the corresponding ETF. Also, larger investors normally will benefit from investing in ETFs, especially if their holding period is of sufficient length. The "sufficient length" depends upon the amount invested. The more invested, the shorter the length of time for the benefits of ETFs to outweigh the costs.

Gastineau (2004) looks at the question of why ETFs underperform index funds that track the same index. His focus is the operational efficiency of the funds' management. By inspecting the historical returns (through 2002) of iShares, Spiders, and Vanguard indexes following the S&P 500 and the Russell 2000, he notes that the ETFs typically underperform their respective index funds. He conjectures that a significant portion of the underperformance is likely due to the failure of ETF fund managers to reduce their transactions costs in a way that is common among index fund managers. When indexes change their composition and/or weighting, the index fund manager will time his modifications to the fund in order to minimize transactions costs. He also notes that although there are no legal barriers against this timing in ETFs, ETF managers have not yet adopted this method of cost reduction. The explanation for the ETF underperformance Gastineau posits seems reasonable, but since there is no analytical methodology incorporated in

his paper, a broader, academic study is clearly needed to determine if his argument has merit. We look at a broader range of ETFs and a more current data set to see if the underperformance holds for the broader range of ETFs or just the small subset used by Gastineau.

In an early work on ETFs, Olienyk et al. (1999) find short-term causal relationships between the ETFs, suggesting arbitrage and opportunities market inefficiencies. Demaine (2002) points out that sector-based ETFs were inevitable, especially in Europe. This is because recent trends have indicated that country-based correlations are increasing, making diversification by country less meaningful in Europe since the introduction of the European Union (Eurozone). However, the industry-based correlations have been declining at the same time, making sector-based investing more attractive. Engle and Sarkar (2002) conclude that domestic ETFs are normally priced very close to their market value. They have less confidence when considering international funds which are less actively traded and therefore less precisely priced. However, they surmise that since these ETFs operate in a more stringent environment, they may still be close to correctly priced. On the other hand, Madura and Ritchie (2004) find a strong overreaction in ETFs during the 1998 – 2002 period when technology issues appear to have been over-priced. They found greater overreaction in ETFs that were more volatile and in international ETFs.[2] Jares and Lavin (2004) find that the daily returns of Asian (specifically Japan and Hong Kong) iShares are well correlated with daily S&P 500 returns. These ETFs show significant power when trying to predict movements in the respective country stock markets. Jares and Lavin (2004) also show that profitable trading strategies may be implemented that take advantage of Asian iShares' overprediction of next day returns to NAV. This line of literature suggests that there is a

[2] Also, using European iShares, Simon and Sternberg (2005) find significant overreaction to after hours developments. So much so that one can derive a profitable trading strategy based upon buying or shorting iShares at the close of U.S. trading, if the discount or premium to NAV is sufficiently large.

fundamental difference between domestic and international ETFs and index funds. Therefore, it may be useful to account for this difference in our analysis.

II.B. Tracking Error

Several authors look at improving (reducing) tracking error through active trading strategies. Jorion (2003) and El-Hassan and Kofman (2003) look at active portfolio allocation strategies that exploit the predictability in the conditional variance-covariance matrix of asset returns by deriving a constrained tracking-error efficient frontier.[3] Then, they use the variance-covariance matrix to rebalance their portfolios to minimize tracking error. In theory, these techniques seem efficient in reducing tracking error. However, in practice, since trades are made dynamically, implementing the strategies would seriously increase the trading volume and turnover ratio in a mutual fund portfolio, thereby increasing trading costs substantially.

Frino and Gallagher (2001) empirically study index fund tracking error. They explain that the primary factor that causes index fund tracking error is the cost of transactions which includes liquidity concerns, fund cash flows, dividends, volatility of the benchmark, corporate activity, and index composition changes. They find that the tracking error associated with the S&P 500 index funds follows a quarterly (seasonal) pattern. The tracking error is lowest at the end of each calendar quarter. They surmise that this seasonal effect may be the result of the timing of

[3] Burmeister et al. (2005) develop diagnostic tools to evaluate alternative active trading strategies for reducing tracking error.

dividend payments by the funds. Other reasons for tracking error, they conclude, may have to do with the changes in the index itself.[4]

III. ANALYSES AND MODELS

The analyses and models are grouped such that the first set deals directly with performance differences between ETFs and index mutual funds and reasons for these differences. The second section of analyses addresses the important issue of tracking error.

III.A. ETF Performance Differences with Index Mutual Funds

ANALYSIS 1: Comparing abnormal returns of exchange-traded funds and mutual funds.

Although they do not specifically compare abnormal returns, Poterba and Shoven (2002) compare total returns of Spiders and the Vanguard Index 500. They show that the Vanguard index mutual fund performed slightly better during the period from 1994 to 2000. Also, Dellva (2001) compares these two funds with the iShares S&P 500 to test which is the best investment. He includes expenses, management fees, and tax considerations in determining his results over one to 15-year time horizons. He concludes that ETFs have advantages over longer holding periods and when larger lump-sum investments are made. Neither study does any kind of regression analysis. Both of these comparisons demonstrate a need for further examination on the topic.

[4] Agapova (2006) also examines tracking error by comparing the tracking error of several ETFs and index mutual funds relative to their indexes. She concludes that ETFs have smaller tracking errors and lower expenses.

To test whether or not ETFs and index mutual funds have insignificantly different abnormal returns, first, we test whether the difference between ETF returns and index fund returns is significantly different using individual comparison tests. We perform Kolmogorov-Smirnov and Shapiro-Wilk tests to examine the data for a normal distribution. Since essentially none of the indexes are normally distributed, we run the nonparametric Wilcoxon Signed-Rank Test on each pair in the data. We test all combinations of index funds and ETFs that follow the same index. If we find a significant result for this test, it will tell us that the pair of funds (or an index and a fund) being tested have different median returns.

The second test of abnormal returns we use is to run OLS regressions of the ETF and index fund returns against the returns of the index that each is tracking. The only independent variable in these equations is the index that the fund is tracking.

$$ETFRet_t = \alpha_{ETF} + \beta_{ETF} INDEX_t + \varepsilon_{ETF} \quad (1)$$

$$IFRet_t = \alpha_{IFUND} + \beta_{IFUND} INDEX_t + \varepsilon_{IFUND} \quad (2)$$

To determine whether the ETF and index fund abnormal returns are significantly different from each other, we then simultaneously run two regressions at a time using a seemingly unrelated regressions (SUR) model and subtract the ETF alpha from the index fund alpha to test whether there are significant differences in the abnormal returns of the paired equations. This test was done for each pair of ETFs and index funds that follow the same index. If we find a significantly

positive result, then we can say for that pair of funds, the index fund performs better than the ETF.

We then perform a third test, a panel regression on all index funds and ETFs, to see whether ETF returns are significantly different from index fund returns:

$$\text{Return}_{p,t} = \alpha + \beta \text{INDEX}_{p,t} + \chi ETF_{p} + \varepsilon_{p,t} \qquad (3)$$

where ETF is a dummy variable that does not vary with time. ETF is equal to one if the fund is an ETF and zero if an index fund. If the dummy coefficient, χ , is significant, then, as a group, index funds and ETFs have significantly different abnormal returns.

Equation (1) also shows something interesting when comparing ETFs that track the same index. There are two cases (shown below) where different ETFs track the same index:

ETF	INDEX TRACKED	ETF ORGANIZATION
SPY (SPDRs - Spiders)	S&P 500	Unit Investment Trust
iShares S&P 500 Index (IVV)	S&P 500	Open-End Investment Company
MDY (MidCap Spiders)	S&P Midcap 400	Unit Investment Trust
iShares S&P MidCap 400 Index (IJH)	S&P Midcap 400	Open-End Investment Company

ANALYSIS 2: Testing ETFs that track the same index.

The Standard & Poor's ETFs were set up as unit investment trusts and are larger and more popular (more trading volume) than the ones set up as open-end investment companies (mutual funds). The S&P ETFs are also at least five years older, and therefore, more established than the

iShares funds which have only been available since May 2000. Again, to test the abnormal returns, we run OLS regressions of the respective ETF returns against the returns of the index each is tracking. To determine whether the ETF and index fund abnormal returns are significantly different from each other, we again simultaneously run regressions with the SUR model and test whether there are significant differences in the alpha terms in the paired equations. If there is a significant difference, it will tell us which ETF outperforms the other.

ANALYSIS 3: Determinants of abnormal returns in ETFs and index funds.

To investigate the determinants of the portfolio's alpha, we run the following regression on alpha to determine if these factors are important in determining the abnormal returns of index funds and ETFs:

$$\hat{\alpha}_i = C_i + \sum_{j=1}^{n} \gamma_{i,j} V_{i,j} + \omega_i D_i + \sum_{j=1}^{n} \lambda_{i,j} V_{i,j} D_i + \eta_i \qquad (4)$$

where $\hat{\alpha}$ is the alpha on the portfolio as determined in equations (1) and (2), C is a linear constant, γ is a slope coefficient, *V* are the variables related to the fund from *j* = 1 to *n* (see below), ω is an adjustment for the constant term given the fund is an ETF, D is a dummy that is equal to one if the fund is an ETF and zero if the fund is an index mutual fund, and λ is a slope adjustment given the fund is an ETF. Some of the variables (*V*) used are inception date, expense ratio, size (total net assets), portfolio risk (standard deviation), turnover ratio, and whether the fund is a no-load fund. Most index funds are no-load funds, so the no-load variable is introduced as a dummy to capture the few index funds that are not no-load. These variables are used because they have been shown in prior literature to affect abnormal returns of mutual funds.

We also use a dummy signifying whether the index fund is a true index fund or not and the relative size (by net assets) of the fund to all funds following the same index. True index funds are defined here to be the index funds that attempt to track the given index by maintaining positions in the assets within the index and trying to replicate the performance of the portfolio on a nearly one-to-one basis. In other words, the relative beta should be very close to one to be considered a true index fund. Index funds that follow an index, but actively target a beta of 2.0 would not be considered a true index fund, for example; nor would a fund that sells the index portfolio short, effectively targeting a beta of –1.0. There are twenty such index funds in our sample. We list them separately as "other index funds." ETFs are all considered true index funds. Relative size is tested to see if fund level market concentration is a factor in determining abnormal returns. All factors described thus far are fund-related and as such, the managers of the funds may have some control over them. If the coefficients on (for example) expense ratio and turnover are negative and significant, then they inversely impact abnormal returns.[5] In this case, the manager would want to minimize expenses and turnover to allow the fund to achieve the highest possible return for investors.

We also introduce two index-specific variables are used that do not change from fund to fund. They are the number of funds following an index, and the Herfindahl index using the assumption that all funds following the same index constitute an industry. These are introduced to see if the concentration of the industry (funds following an index) has an effect on abnormal returns. We assume that the fund manager has no control over this factor.

[5] Commissions are not included because they typically are fixed and mainly vary by broker rather than strictly by shares or dollars transacted.

III.B. Tracking Error

Many studies examining ETFs and index funds make a point of determining each fund's tracking error. When reading the literature, one might wonder why investors would care about tracking error. One reason to look at tracking error is because investors use ETFs to change equity positions quickly. For example, hedge funds use ETFs frequently. Tracking errors are also important to arbitrageurs, who could profit from misalignment of fund prices and the underlying index. By observing ETFs, we may be able to gauge what hedge funds are doing in a real time basis. If ETFs show a poor tracking ability, they would be less useful for hedge funds. Therefore, we analyze ETFs and index funds to compare their abilities to track the indexes they follow.

ANALYSIS 4: Comparison of ETF and index mutual fund tracking errors.

When analyzing the tracking error of ETFs and index funds, we look at the funds in a manner similar to that used by Frino and Gallagher (2001). Using 42 index funds that track the S&P 500, Frino and Gallagher calculate tracking error (TE) in three ways. We will use the same notation they use.

In the first method, we derive tracking error by using the absolute value of the difference in returns of the ETF or index fund and the benchmark index ($e_{pt} = R_{pt} - R_{bt}$), starting on day t, where the daily average absolute tracking error over n days ($TE_{1,p}$) is:

$$TE_{1,p} = \frac{\sum_{t=1}^{n} \left| e_{pt} \right|}{n} \tag{5}$$

A second test for tracking error compares the variability (standard deviation) of the difference in the ETF or index portfolio returns and the underlying index return ($TE_{2,p}$). This method may be calculated as:

$$TE_{2,p} = \sqrt{\frac{1}{n-1} \sum_{t=1}^{n} \left(e_{pt} - \bar{e}_p \right)^2} \tag{6}$$

The third method used for tracking error is to estimate the standard error of the residuals of a regression on returns ($TE_{3,p}$). The model is simply the same as models (1) and (2). This method is useful because when regressing the return of the ETF or index portfolio on the return of the underlying index, the standard error of the regression equation provides a general estimate of tracking error. Pope and Yadav (1994) explain, however, that if the beta of this regression is not exactly equal to one, the regression residuals will be biased, and therefore, differ from $TE_{2,p}$. These authors show that due to negative serial correlation, tracking error will be overstated if the portfolio does not have a beta of exactly 1.0 against the benchmark portfolio. Negative serial correlation is especially a problem when using higher frequency data such as daily or weekly returns. However, the emphasis of this study is the difference in the tracking errors between the two types of fund. So, since there is no evidence that these measurement problems are more serious for one type of fund, we do not expect that the negative serial correlation will lead to biased conclusions.

HYPOTHESIS: Tracking error has an inverse relationship with fund flows

After calculating tracking error, we regressed lagged tracking errors on fund flows of ETFs and index funds to determine if the tracking error affects fund flows. The purpose of this test is to gauge the importance of tracking errors to investors. If investors view tracking errors as important for trading, they will move money in and out these funds based upon tracking errors.

$$FF_{t,p} = \psi_{t,p} + \kappa_{t,p} TE^{n}_{t-s,p} + \sum_{i=1}^{3} \tau_{i,t,p} \delta_i + \sum_{i=1}^{3} \theta_{i,t,p} \delta_i TE^{n}_{t-s,p} \qquad (7)$$

where FF_t is fund flow at time *t*; ψ, κ, τ and θ are parameters; TE is the tracking error variable at time *t–s*, where *s* = 1, 2; and *n* is the superscript denoting each method of calculating tracking error. There are three dummy variables for this regression. δ_1 is a dummy that equals one if the fund is an ETF and zero if an index fund; δ_2 is a dummy that equals one if the fund is an international fund and zero if it is not an international fund; and, δ_3 is a dummy that equals one if the fund is a true index fund and zero if it is not a true index fund. We use the second dummy because there is some evidence that international funds are less liquid and less well priced (Engle and Sarkar, 2002). Therefore, the tracking errors of international funds are higher than domestic funds. The third dummy is used because of the large differences found in index funds that do not track their indexes targeting a beta of 1.0. We look at one and two lagged periods for tracking error to see if past tracking error affects fund flows further away in time. If κ_p is significant, then past tracking error affects current fund flows. Our a priori expectation is

that κ_p will be negative and significant, which means if tracking error is large, fund flows will decrease by more than if tracking error is small. Since fund flow is significantly impacted by hedging and arbitraging activities, the significance of κ also would indirectly suggest that tracking error is important to participants in these activities. The κ_p obtained from the differing tracking error methods should be very similar. We also expect the intercept term, $\psi_{t,p}$, to be positive and significant implying the flow of funds is typically positive into both ETFs and index mutual funds. Also given the popularity of ETFs, they are generally growing at a faster pace.[6] So, the intercept may be greater for ETFs than for mutual funds.

IV. DATA

The data in this paper come from various sources. General ETF data except for prices and returns come from "Mutual Funds: The Individual Investor's Guide to Exchange-Traded Funds 2005" which is an annual article in the trade journal sponsored by The American Association of Individual Investors. ETF daily prices, returns and other supplemental ETF data are from Yahoo Finance.

Index mutual fund data except for prices and returns come from Morningstar Principia Advanced Mutual Funds Module dated February 29, 2004. From this CD, we obtained all of the data for the index fund characteristics such as net asset value (size), turnover and expense ratio. Daily index fund prices and returns are from Yahoo Finance.

[6] Agapova (2006) finds that flows into ETFs were positive and substantially higher than flows into index mutual funds.

Calvert, Morningstar Indexes, Morgan Stanley Capital International (MSCI), and Yahoo Finance provided index values. Funds less than two years old on June 30, 2006 are excluded from the data set because two years is the minimum amount of data required to determine the annual fund flows. As much as possible, ETFs are matched with index funds following the same index. Fund flow data are annual and are obtained directly from fund prospectuses.

The final data set includes 177 index funds and 72 ETFs that follow 78 indexes. Of these, there are 25 indexes that have at least one index fund *and* one ETF that follow it (several have multiple index funds following the index). A total of 27 funds, two index funds and 25 ETFs, are international funds, and 20 index funds were not classified as "true index funds."

V. EMPIRCAL RESULTS

In order to determine whether parametric or nonparametric tests were appropriate for our paired comparison tests, we first perform Kolmogorov-Smirnov and Shapiro-Wilk tests to determine whether each (or any) of the indexes follow a normal distribution. In seven out of the 78 indexes, a normal distribution could not be ruled out. All indexes for which we were unable to rule out a normal distribution have existed for less than three years.

Because the indexes do not follow a normal distribution, a paired t-test to determine whether index funds and ETFs have similar distributions to their respective indexes is not entirely appropriate. Therefore we performed a nonparametric test, the Wilcoxon Signed-Ranks Test, on

the daily returns for each of the following pair-types following the same index: index/index fund, index/ETF and index fund/ETF. A test with a significant p-value indicates the differences in the median returns of the pairs are significantly different. In other words, the index, index fund or ETF does a poor job of reproducing the same median return as its counterpart.

Looking at the Wilcoxon Signed-Ranks Test results in Table 2, it is apparent that a large number of index funds do not produce the same median return as their indexes (58% of the funds had significant p-values at the 10% level or lower). We also note that of the 102 statistically significant comparisons, the index outperforms the index fund 87% of the time. This result is expected since expenses reduce the funds' returns. However, when examining the other three panels, we find only a small number of other pair-types do a poor job of reproducing the returns of their paired asset. We expected ETF-index fund and ETF-ETF pair-types to match fairly closely because both ETFs and index funds have both been shown to underperform their benchmarks[7] so ETFs and index funds should (or at least could) have very similar median returns, but continue to underperform relative to the index. However, the most unexpected result from this test is the Index-ETF test comparisons. These test results suggest ETFs *do* produce the same median returns as their indexes. Only 7% of these comparisons are significant at the 10% level or below and none is significant at the 1% level.

We summarize the results from the individual OLS regressions on index funds and ETFs in Table 3. We find that the alphas (abnormal returns) are very small using our daily data. There are 256 index-fund-ETF pairs in our sample that each follow the same index. Descriptive results from the regressions are shown in Panel A. The range of differences in alphas is between

[7] See Elton, Gruber and Busse (2004) for index funds; Gastineau (2004) and Engle and Sarkar (2002) for ETFs.

+0.0321% and –0.0245% per day. Only nine of the 256 pairs have differences in alphas which are statistically significant. These are individually listed in Panel B of the table. Even though the results were generally insignificant, there some results from these regressions that are noteworthy. First, out of the 256 comparisons, the abnormal returns associated with the ETFs are higher than the alphas of the index funds in 214 cases, which represents 83.6% of all cases. This result is in contrast to previous results that conclude that index funds tend to have higher alphas than ETFs. This result may be due to ETF managers reducing their costs by timing their modifications to the fund in order to minimize transactions costs, especially in funds that follow the S&P 500. The results are much more prevalent in funds that follow the S&P 500 than funds that do not. About 86% of S&P 500 funds have higher performing ETFs, while only about 30% of ETFs that follow other indexes outperform their index fund counterparts. This result is not altogether unexpected since S&P 500 funds are generally more scrutinized than other funds. Another result of note from this table is that of the nine significant results, only three indexes were involved and most of the significant results were between ETFs and index funds that were not classified as "true" index funds. Overall, this table suggests that individually, index funds typically do not significantly differ in performance with their ETF counterparts. However, as a group, the mean abnormal return for ETFs is statistically significantly higher than that of index funds.

After analyzing the individual regressions in Table 3, Table 4 shows the results of the pooled regressions. All returns from index funds and ETFs were pooled along with their respective indexes. A dummy was added to the equation to extract the differences between index funds and ETFs as shown in equation (3). The results show that when all returns are regressed against their

respective index returns, the ETF dummy is weakly significant. So, to delve into the reason for the significant results, we included three more regressions. When international funds are removed, the significance disappears, but when only true index funds and ETFs are regressed, the ETF dummy increases its significance. As mentioned earlier, true funds are ones that attempt to track the given index by maintaining positions in the assets within the index and trying to replicate the performance of the portfolio on a one-to-one basis. All ETFs fall into this category, as do all but 20 index funds. In the last regression in this table, when both international funds and non-true index funds are excluded, the ETF dummy is no longer significant. Essentially, this table is confirming results from the previous table, the index funds that were categorized as not being true index funds seem to be driving the differences between ETFs and index funds. So, index funds which target an exact replication of their index's performance do not significantly differ in performance with their ETF counterparts. It is also noteworthy that the ETF dummy is always positive, again confirming the findings in the previous table.

Next, we look at the elements that may affect the abnormal returns. The factors that we include as possible determinants are inception date, expense ratio, net assets, standard deviation, turnover, and whether or not the fund is a no load fund. These variables are used because they have been shown in prior literature to affect abnormal returns of mutual funds. We also include as possible factors that influence abnormal returns are whether the fund is a true index fund (dummy), and the relative size of the fund when compared to all funds following the same index. For this variable, if a fund is the only fund following a given index, its relative size under this measure would be 100%. We also use two index-specific variables that do not change from fund to fund. First is the number of funds following an index, and the second is the Herfindahl index

using the assumption that all funds following the same index constitute an industry. If only one fund follows a given index, its Herfindahl index would be 10,000. We include these factors to see if fund-level or industry-level concentration is important in determining the abnormal return of the funds.

Some variables are relatively highly correlated with other variables. The correlations between the independent variables are shown in Table 5. Some of the very high correlations are expected. For example, RelSize (relative size), FundsIdx (funds following an index), and HerfIdx (Herfindahl index) are all trying to capture the same idea, specifically, how the concentration of the market affects abnormal return. The ETF dummy is correlated with all three of these variables since ETFs generally have higher concentration percentages than index funds in our sample. This higher concentration may be attributed to the fact that usually only one ETF follows a given index whereas often several mutual funds will follow the index. So all money going into ETFs following an index gets funneled into a single fund, but there are several choices for money going into index funds. The inception date is also well correlated with FndsIdx and HerfIdx, suggesting that indexes with newer funds following it tend to have a smaller number of funds, and higher concentration among funds following the index. TrueIndex, the dummy variable that describes whether or not the fund was a true index fund or not, is highly correlated with expense ratio, standard deviation, and turnover, implying that funds targeting a beta of 1.0 with their indexes have lower expenses, lower standard deviation, and lower turnover, all of which one would expect since the other funds all target absolute betas higher than 1.0 (i.e. higher than 1.0 or lower than −1.0) and usually have higher volatility in their returns. The most difficult correlation to understand is that expense ratio is positively correlated with abnormal returns at

the 5% significance level. This implies that higher expenses are correlated with higher abnormal returns, which is counter-intuitive. However, when one realizes that some of the correlations may be driven by the funds that are not true index funds, this correlation makes more sense. The funds not classified as true index funds have alphas approximately 73% higher than true index funds. Other unusual differences are expense ratio, net assets, standard deviation, and turnover, which are much different than true index funds. Expense ratios are more than three times higher; net assets are (on average) less than one-tenth, standard deviation is more than double, and turnover is more than 13 times higher than true index funds. When removing these funds, the abnormal return and expense ratio are no longer significantly correlated.

Table 6 illustrates the results of these regressions using equation (4). Two variables, the no load dummy (NL) and the true index dummy (TrueIndex) were not multiplied by the ETF dummy and included in the regressions since they are only applicable to index funds. For this table, each of the factors is listed on the left and seven regressions are shown that include the standard variables found in prior literature to be important determinants of abnormal return along with other variables that were considered for inclusion in the regression described above. All combinations were tested. However, only the most relevant are shown due to space considerations. Each of the four variables, TrueIndex, RelSize, FundsIdx, and HerfIdx were included one at time to be sure that they added something to the regression. TrueIndex made the regression worse. In all cases, including TrueIndex reduced the adjusted R-squared. Therefore, it was not considered for the final equation. Since FundsIdx, and HerfIdx are highly correlated and are included for the same purpose, to measure the industry concentration, we only want to include one of them in a final regression. RelSize is highly correlated with these two variables as

well, but it is a fund-specific variable, while FundsIdx and HerfIdx are industry-specific, so it was still considered in regressions that also include either FundsIdx or HerfIdx.

The Herfindahl index always helped our regressions slightly more than the number of funds following an index. So, we conclude the best regression is in the last column of Table 6. Therefore, we find the ETF dummy is significant in determining abnormal return suggesting ETFs have a significantly higher abnormal return than index funds for our sample. Also, for index funds, the relative size and Herfindahl index are significant determinants, while for ETFs, the inception date, expense ratio, standard deviation, and relative size are all important determinants of abnormal return. These findings suggest that managers may be able to improve performance by reducing the expense ratios and standard deviations of their portfolios. If they have funds competing with them by tracking the same index, the fund managers may also consider limiting the total net assets allowed to be invested into the funds since funds with lower relative net assets tend to perform better than larger ones.

Next, we begin our look at tracking error by analyzing the correlations of all 507 pair-types. The correlations are important to determine if there are any significant differences in the funds that should be accounted for in later analyses. We chose the Spearman correlation since the data is not normally distributed. Table 7 shows the results of the Spearman correlations. The primary finding is that almost all funds that are not following international indexes have very high correlations with their indexes. Over 98% of the pairs have absolute correlations higher than 90%. In fact, almost one-quarter of the pairs have correlations above 99%. We are not surprised that most funds have a high correlation with the index that the fund is following. After all,

having a high correlation with the index is often the primary goal of the ETF or index fund. The high Spearman correlations suggest most of the funds track their indexes well and are well-correlated with other funds tracking the same index.

Index funds and ETFs following international indexes reveal quite a different story, though. None of these funds is correlated with its index at the 95% level, and most are only correlated between 50% and 70% with their indexes. Only one international pair, the MSCI Pacific Excluding Japan (index) and iShares MSCI Pacific Ex-Japan Index Fund (ETF), had a lower correlation than 50%. This result may be due to the previous finding that country funds tend to be more correlated with domestic stock indexes in the short run, while being more correlated with the respective foreign index in the long run (see Chiang and Kim, 2003). In general, from the correlations we find that we must account for international funds separately to be sure that their poor correlations do not affect our results in an unforeseen way.

Table 8 shows some descriptive results for tracking error. Results are broken down into ETFs and Index Funds, and further into Domestic and International ETFs and "True Index Funds" and Other Index Funds. When examining the general results for ETFs and index funds, it appears that index funds are able to track their indexes much better than ETFs. After filtering out the international ETFs and the "other index funds," the results show that domestic ETFs are much better at tracking their indexes than their international counterparts. In the same light, true index funds are significantly better at tracking their indexes than the other index funds. When comparing the domestic ETFs to the true index funds, the index funds still track the indexes better as a group. However, the index funds that were not true index funds were the worst group

at tracking the index, even worse than the international ETFs. This table further confirms the differences in domestic and international funds as well as the differences in true index funds and other funds.

In Table 9, results analyzing whether tracking error is a significant factor in the flow of funds for index funds and ETFs is presented. We show heteroskedasticity-consistent regressions for one- and two-period fund flows using each of the three different tracking error methods. There are a few notable results from these regressions. For one-period fund flows in Panel A, our a priori expectation that the constant is positive and significant is confirmed. This implies that there is a positive inflow of funds into index funds and ETFs irrespective of other factors. Another interesting result from this table is that tracking error *does* affect fund flow in the following period. Greater tracking error is associated with a decrease in fund flow during the following period, as we expected. The adjustment for ETFs is not significant, though. An additional finding from this table is that the third tracking error method, the standard error of the regressions, seems to be the weakest. We expected this as well because of the inclusion of funds in the regression that are not "true index funds." As stated earlier, Pope and Yadav (1994) show that this third tracking error method is biased, and therefore not reliable, unless the beta of the regression is exactly equal to one. The first two tracking error methods are nearly equal, with the second tracking error method yielding only slightly stronger results than the first method.

In Panel B of Table 9, a regression for the second year is performed to see if results carried over to the second year. Almost none of the results for the first year carry forward into the second year. The only significant result here is the importance of the international funds. The apparent

suggestion of this result is that international funds have a significant reduction in fund flow in the second year due to greater tracking error. We can see no reasonable explanation for this other than that the result may have been due to an unknown unique characteristic of international funds. Investigation into this international phenomenon may be worth pursuing in a future study. So our conclusion from this table is that tracking error negatively affects fund flow in the following period, but is not a factor beyond one period. This result shows that investors pay attention to tracking error when deciding upon which investment to make. If funds are not very good at tracking the index that they are following, investors will put more money into other funds following the same index that track the index better.

VI. CONCLUSION

Using daily data, we find that no indexes that have existed for more than three years have returns that follow a normal distribution. Therefore, we use the nonparametric Wilcoxon Signed-Ranks Test and show that a large number of index funds do not reproduce the same median returns as their indexes. However, when ETFs are compared with their indexes, there are generally no significant differences. When we compare index funds to ETFs following the same index, we find no significant differences between the medians of the returns, suggesting that neither index funds nor ETFs can significantly outperform the other.

By regressing index funds and ETFs on the indexes that they follow, we find the abnormal returns are very small and virtually none were significant. However, the abnormal returns associated with the ETFs are higher than the alphas of the index funds in 83.6% of all cases. This

result is confirmed by a pooled regression analysis and is in contrast to previous research which concludes that index funds tend to have higher alphas than ETFs. This new finding may be due to our use of a much broader array of ETF in this study while previous studies are much more limited in scope, generally examining one or two ETFs. We also note that the results are much more prevalent in funds that follow the S&P 500 than funds that do not. Only about 30% of ETFs that follow other indexes outperform their index fund counterparts. This finding may be the result of S&P 500 ETF managers improving their performance by timing their modifications to these ETF so that they minimize transactions costs, a way Gastineau (2004) says is common among index fund managers. Essentially, our results suggest that if an investor wants to track the S&P 500, he or she should invest in ETFs, but if investing to track another index, it would be better to invest in index funds.

When looking at the components of abnormal returns, several regressions were performed. We find again that ETFs have higher abnormal returns than index funds. Other important factors for index funds include the relative size of a fund when compared with other funds following the same index and the Herfindahl index (given funds following a single index constitute an industry). For ETFs, the inception date, expense ratio, standard deviation, and relative size are all important determinants of abnormal return. These results suggest that managers may have some control over the performance of ETFs by (for example) reducing expenses and the standard deviation (risk) of their investments. However, it is not as simple for managers of index funds. The factors that are important are related to the industry level market concentration of the funds following each index. To improve performance, index fund managers would need to be able to control not only their funds, but other funds that follow the same index.

When examining the tracking errors for ETFs and index funds, we first perform Spearman correlations on each of the 507 pair-types. They show that almost all funds not following international indexes have very high correlations with their indexes. However, index funds and ETFs following international indexes are only correlated with their indexes between 50% and 70% with their indexes. Also, we generally find index funds are able to track their indexes more closely than ETFs. After filtering out the international ETFs and the "other index funds," the results show that domestic ETFs are much better at tracking their indexes than their international counterparts. Also, true index funds, defined as index funds that attempt to track their respective indexes by trying to replicate the performance of the index portfolios with a relative beta of one, are significantly better at tracking their indexes than the other index funds. When comparing the domestic ETFs to the true index funds, the index funds still track the indexes better as a group. However, the index funds that were not true index funds were the worst group at tracking the index, even worse than the international ETFs. These results inform us that when we are looking at fund flow in the next equation, we must account for international funds and funds that are not true index funds separately.

Finally, we analyzed tracking error to see if it is a significant factor in the flow of funds for index funds and ETFs. We find that tracking error *does* affect fund flow for one period forward, but not two. As predicted, greater tracking error is associated with a decrease in fund flow. Also we find that there is a positive inflow of funds into index funds and ETFs irrespective of other factors. An additional finding from this analysis is that the standard error of the regressions seems to be the weakest tracking error method. These results suggest that those who invest in index funds

and ETFs monitor tracking error and invest in them based at least partially based upon the ability of the fund to closely track the index. The closer a fund tracks its index, the more likely the investor is to invest additional funds.

REFERENCES

Agapova, Anna, "Innovations in Financial Products: Conventional Mutual Funds versus Exchange Traded Funds," August 2006, Working Paper, Georgia State University.

Chiang, Thomas C. and Doseong Kim, "On Country-Fund Price Behavior – An Empirical Analysis of Co-Integrating Factors," *Advances in Financial Planning and Forecasting*, 2003, vol. 11, pp.85 – 112.

Dellva, Wilfred L., "Exchange-Traded Funds Not for Everyone," *Journal of Financial Planning*, 2001, April, vol. 14, no. 4, pp. 110 – 124.

Demaine, John, "Exchange Traded Funds for the Sophisticated Investor," *Derivatives Use, Trading & Regulation*, 2002, vol. 7, no. 4, pp. 354 – 361.

El-Hassan, Nadima and Paul Kofman, "Tracking Error and Active Portfolio Management," *Australian Journal of Management*, 2003, vol. 28, no. 2, pp. 183 – 207.

Elton, Edwin J., Martin J. Gruber, George Comer and Kai Li, "Spiders: Where Are the Bugs?," *Journal of Business*, 2002, vol. 75, no. 3, pp. 453 – 472.

Engle, Robert and Debojyoti Sarkar, "Are Exchange Traded Funds Fairly Priced?," October 2002, Working paper, New York University.

Frino, Alex and David R Gallagher, "Tracking S&P 500 Index Funds: Getting Your Money's Worth?," *Journal of Portfolio Management*, 2001, Fall, vol. 28, no. 1, pp. 44 – 55.

Gastineau, Gary L., "The Benchmark Index ETF Performance Problem: A Simple Solution," *Journal of Portfolio Management*, 2004, Winter, vol. 30, no. 2, pp. 96 – 103.

Jares, Timothy E., and Angeline M. Lavin, "Japan and Hong Kong Exchange-Traded Funds (ETFs): Discounts, Returns and Trading Strategies," *Journal of Financial Services Research*, 2004, vol. 25, no. 1, pp. 57 – 69.

Jorion, Phillipe, "Portfolio Optimization with Tracking-Error Constraints," *Financial Analysts Journal*, 2003, vol. 59, no. 5, pp. 70 – 82.

Kostovetsky, Leonard, "Index Mutual Funds and Exchange-Traded Funds," *Journal of Portfolio Management*, 2003, Summer, vol. 29, no. 4, pp. 80 – 92.

Madura, Jeff and Nivine Richie, "Overreaction of Exchange-Traded Funds During the Bubble of 1998-2002," *The Journal of Behavioral Finance*, 2004, vol. 5, no. 2, pp. 91 – 104.

Olienyk, John P., Robert G. Schwebach and J. Kenton Zumwalt, "WEBS, SPDRs, and country funds: an analysis of international cointegration," *Journal of Multinational Financial Management*, 1999, vol. 9, pp. 217 – 232.

Pope, Peter F., and Pradeep K. Yadav, "Discovering Errors in Tracking Error," *The Journal of Portfolio Management*, 1994, vol. 20, no. 2, pp. 27 – 32.

Poterba, James, M. and John B. Shoven, "Exchange-Traded Funds: A New Investment Option for Taxable Investors," *American Economic Review*, 2002, May, vol. 92, no. 2, pp. 422 – 427.

Table 1. ETF-Index Fund Cost Comparison

Fund Costs

Types of Costs	Exchange-Traded Funds	Traditional Index Funds
Fund Transaction Costs	None (redemption-in-kind)	Bid-Ask spreads to market makers
Cash Inflows and Outflows	Deviations in value of creations and redemption in-kind	Cash drag for uninvested funds
Dividend Policy	Lag between ex-dividend date and receipt of dividends	Lag between ex-dividend date and receipt of dividends
In-and-Out Arbitrage Trading	None due to the way shares are created and redeemed	Can be important for some domestic index funds.
Index Fund Changes	Rebalance transactions costs	Rebalance transactions costs
Corporate Activity	Rebalance transactions costs	Rebalance transactions costs
Management Fees	Similar to mutual funds, but shareholder accounting is done by the shareholder (less cost)	Similar to ETFs, but shareholder accounting is done by the fund (more cost)

Shareholder Costs

Types of Costs	Exchange-Traded Funds	Traditional Index Funds
Shareholder Transaction Costs	Broker commissions + bid-ask spreads	None, except for index funds with loads
Taxation Costs	Virtually no capital gains distributions or taxation costs	Significant share of capital gains are distributed
Exchange Fees	None, except for transaction costs	Charged when an investor transfers money between funds within the same fund family
Account Maintenance Fee	Charged by some brokerages to maintain certain accounts	Charged by some funds to maintain certain accounts
12b-1 Fees	Not applicable to ETFs	Fee charged by some funds to compensate for marketing costs

Table 2. Wilcoxon Sign-Rank Tests

PANEL A. Index to Index Fund

n = 177

P-level Significance	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1%	78	44.07%	78	44.07%
5%	11	6.21%	89	50.28%
10%	13	7.34%	102	57.63%

PANEL B. Index to ETF

n = 72

P-level Significance	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1%	0	0.00%	0	0.00%
5%	2	2.78%	2	2.78%
10%	3	4.17%	5	6.94%

PANEL C. ETF to Index Fund (Following same index)

n = 256

P-level Significance	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1%	1	0.39%	1	0.39%
5%	2	0.78%	3	1.17%
10%	1	0.39%	4	1.56%

PANEL D. ETF to ETF (Following same index)

n = 2

P-level Significance	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1%	0	0.00%	0	0.00%
5%	0	0.00%	0	0.00%
10%	0	0.00%	0	0.00%

We perform a Wilcoxon Signed-Rank Test on 507 combinations of indexes, index funds, and ETFs. Each panel above illustrates a specific pair-type. The column "Frequency" illustrates the number of tests of that pair-type that fall between the specified significance levels (i.e. for Panel A, 11 pairs have significance levels (p-values) falling between 1% and 5%).

Table 3. Index Fund and ETF Abnormal Return Comparisons (Individual Regressions)

PANEL A.

Statistic	Alpha (E)	Alpha (I)	Alpha (I) - Alpha (E)	P-Value for Difference
Mean	0.0054%	0.0023%	-0.0031%	0.0000
Median	0.0061%	0.0033%	-0.0023%	
Maximum	0.0242%	0.0367%	0.0321%	
Minimum	-0.0037%	-0.0250%	-0.0245%	
Standard Dev.	3.164E-05	6.785E-05	6.528E-05	

PANEL B.

Index	ETF	Alpha (E)	Index Fund	True Index = 1	Alpha (I)	Alpha (I) - Alpha (E)	p-value
DJIA	DIA	9.23E-05	RYCVX	0	-4.57E-05	-1.38E-04	0.01864
DJIA	DIA	7.89E-05	PDOWX	0	-5.35E-05	-1.32E-04	0.02317
S&P 500	IVV	6.63E-05	VADBX	1	0.0003668	3.01E-04	0.00417
S&P 500	IVV	6.63E-05	ULPIX	0	-9.81E-05	-1.64E-04	0.00795
S&P 500	IVV	6.42E-05	RYTNX	0	-0.0001349	-1.99E-04	0.01699
S&P 500	IVV	6.63E-05	BLPIX	1	-4.98E-05	-1.16E-04	0.05472
S&P 500	SPY	6.22E-05	ULPIX	0	-0.0001464	-2.09E-04	0.00169
S&P 500	SPY	6.25E-05	RYTNX	0	-0.0001349	-1.97E-04	0.02918
S&P 500	SPY	5.97E-05	BLPIX	1	-5.95E-05	-1.19E-04	0.06461

256 ETF-Index Fund pairs were regressed using $ETFRet_t = \alpha_{ETF} + \beta_{ETF} INDEX_t + \varepsilon_{ETF}$ or $IFRet_t = \alpha_{IFUND} + \beta_{IFUND} INDEX_t + \varepsilon_{IFUND}$ then the alphas were compared using a SUR analysis. Panel A shows the descriptive statistics for these comparisons. Alpha (E) is the abnormal return on the ETF, and Alpha (I) is the abnormal return for the index fund. The fourth column shows the descriptives of the *differences* between the abnormal returns on the index fund and ETF. Panel B shows all nine pairs that have a significant difference.

Table 4. Index Fund and ETF Abnormal Return Comparisons (Pooled Regressions)

Variable	All Funds	Domestic Funds	True Funds	Domestic & True Funds
Alpha	6.91E-05 ***	6.91E-05 ***	5.68E-05 ***	5.40E-05 ***
	1.86E-05	***1.86E-05***	***1.10E-05***	***9.70E-06***
Return on Index	0.80296 ***	0.80013 ***	0.89984 ***	0.9078 ***
	1.32E-03	***1.38E-03***	***7.88E-04***	***7.34E-04***
ETF Dummy	7.29E-05 *	1.79E-05	4.63E-05 **	1.97E-06
	4.06E-05	***4.89E-05***	***2.30E-05***	***2.44E-05***
Adjusted R-Squared	0.48659	0.48267	0.78364	0.82324

Pooled regressions were performed on all index funds and ETFs using the equation, $Return_{p,t} = \alpha + INDEX_{p,t}\beta + ETF_p + \varepsilon_{p,t}$. The first regression shows the results of all 249 index funds and ETFs. The other regressions apply the same equation but remove the 27 international funds in the second regression, remove the 20 funds that are not classified as true index funds in the third regression, and remove both the international funds and the funds that are not classified as true index funds in the last regression. Standard errors are below the coefficients. *, **, *** represent the 10%, 5%, and 1% levels of significance, respectively.

Table 5. Alpha Components Correlation Matrix

Variable	Alpha	ETFDum	Incept	ExpRatio	NetAssets	SDev	Turnover
ETFDum	0.2332 ***						
Incept	0.0018	0.2720 ***					
ExpRatio	0.1461 **	-0.2714 ***	0.0968				
NetAssets	-0.0283	0.1133 *	-0.3250 ***	-0.2057 ***			
SDev	0.2119 ***	-0.0681	-0.0152	0.4747 ***	-0.0525		
Turnover	0.1019	-0.1560 **	0.0762	0.4368 ***	-0.0787	0.3984 ***	
NL	-0.0094	0.2131 ***	0.0855	-0.3376 ***	0.0904	0.0335	0.0747
TrueIndex	-0.1478 **	0.1885 ***	-0.1128 *	-0.6463 ***	0.0823	-0.6111 ***	-0.5922 ***
RelSize	0.2919 ***	0.7781 ***	0.1001	-0.2186 ***	0.1141 *	-0.0287	-0.1558 **
FundsIdx	-0.2658 ***	-0.4778 ***	-0.3469 ***	-0.0250	0.0948	-0.1060 *	0.0494
HerfIdx	0.2369 ***	0.6044 ***	0.3121 ***	0.0264	-0.0614	0.1702 ***	-0.0239

Alpha Components Correlation Matrix Continued

Variable	NL	TrueIndex	RelSize	FundsIdx
TrueIndex	-0.0987			
RelSize	0.1430 **	0.1765 ***		
FundsIdx	-0.0922	0.0229	-0.5739 ***	
HerfIdx	0.1411 **	-0.0706	0.7201 ***	-0.8415 ***

Correlation coefficients are shown for each of the variables that is used in the determination of the components of alpha. ETFDum is a dummy equal to one if the fund is an ETF and zero if an index fund. Incept, ExpRatio, NetAssets, SDev, are the inception date (a higher number implies the fund is newer), the fund's expense ratio, total net assets, and standard deviation of returns. Turnover, NL, TrueIndex, and RelSize are turnover as defined by the fund's purchases divided by average total net assets for the year, dummy indicating whether the fund is a no-load fund or not, a dummy indicating whether the fund is attempting to match its index with a beta of 1.0 (true index) or some other beta, and the size of the fund relative to other funds following the same index. FundsIdx and HerfIdx are the same for all funds that follow a single index. FundsIdx indicates the number of funds following the index, while HerfIdx is the Herfindahl index assuming all funds following a single index represent an industry. All annual values are calculated as of June 30, 2006. *, **, *** represent correlations that are significant at the 10%, 5%, and 1% levels of significance, respectively.

Table 6. Regressions of Alpha Components

Factor	Base Model	TrueIndex	RelSize	FundsIdx	HerfIdx	RelSize&FundsIdx	Relsize&HerfIdx
Constant	3.337E-04	2.995E-04	2.639E-04	4.373E-04 **	3.256E-04	2.887E-04	1.495E-04
	2.108E-04	*2.327E-04*	*2.088E-04*	*2.210E-04*	*2.130E-04*	*2.285E-04*	*2.146E-04*
ETF Dummy	1.762E-03 **	1.770E-03 **	1.747E-03 **	1.607E-03 **	1.703E-03 **	1.750E-03 **	1.846E-03 **
	7.747E-04	*7.765E-04*	*7.631E-04*	*7.872E-04*	*7.754E-04*	*7.804E-04*	*7.604E-04*
Inception	-5.782E-09	-5.738E-09	-4.353E-09	-7.949E-09	-5.509E-09	-4.869E-09	-9.669E-10
	5.940E-09	*5.953E-09*	*5.862E-09*	*6.097E-09*	*6.030E-09*	*6.183E-09*	*6.046E-09*
Inception*ETFDummy	-4.504E-08 **	-4.508E-08 **	-4.258E-08 **	-4.158E-08 **	-4.024E-08 *	-4.264E-08 **	-4.455E-08 **
	2.043E-08	*2.047E-08*	*2.022E-08*	*2.071E-08*	*2.068E-08*	*2.054E-08*	*2.027E-08*
Expense Ratio	-3.373E-03	-2.785E-03	-2.763E-03	-3.661E-03	-3.319E-03	-2.856E-03	-1.933E-03
	2.734E-03	*3.212E-03*	*2.697E-03*	*2.737E-03*	*2.739E-03*	*2.729E-03*	*2.709E-03*
Expense Ratio*ETFDummy	4.125E-02 ***	4.066E-02 ***	4.770E-02 ***	4.251E-02 ***	4.886E-02 ***	4.758E-02 ***	4.860E-02 ***
	1.109E-02	*1.124E-02*	*1.166E-02*	*1.134E-02*	*1.207E-02*	*1.175E-02*	*1.188E-02*
Net Assets	-1.469E-16	-9.982E-17	-8.518E-16	-2.573E-16	-1.222E-16	-8.335E-16	-9.691E-16
	1.719E-15	*1.728E-15*	*1.710E-15*	*1.719E-15*	*1.720E-15*	*1.718E-15*	*1.700E-15*
Net Assets*ETFDummy	-6.169E-16	-6.640E-16	-7.749E-16	-1.354E-15	-2.176E-15	-3.288E-16	-1.193E-15
	2.875E-15	*2.884E-15*	*2.883E-15*	*3.566E-15*	*3.029E-15*	*3.543E-15*	*2.977E-15*
Standard Deviation	-5.327E-04	-1.300E-04	-6.415E-04	-9.565E-04	-4.175E-04	-7.202E-04	4.254E-04
	1.957E-03	*2.274E-03*	*1.924E-03*	*1.974E-03*	*2.007E-03*	*1.954E-03*	*1.979E-03*
Standard Deviation*ETFDummy	-1.387E-02 **	-1.427E-02 **	-1.257E-02 **	-1.341E-02 **	-1.432E-02 **	-1.246E-02 **	-1.423E-02 **
	6.227E-03	*6.344E-03*	*6.160E-03*	*6.226E-03*	*6.239E-03*	*6.194E-03*	*6.208E-03*
Turnover	9.444E-06	1.011E-05	9.348E-06	1.037E-05	9.330E-06	9.539E-06	8.195E-06
	6.308E-06	*6.600E-06*	*6.200E-06*	*6.329E-06*	*6.315E-06*	*6.264E-06*	*6.187E-06*
Turnover*ETFDummy	3.697E-05	3.630E-05	-4.379E-05	3.301E-05	2.379E-05	-4.636E-05	-2.213E-05
	1.739E-04	*1.743E-04*	*1.774E-04*	*1.739E-04*	*1.739E-04*	*1.784E-04*	*1.790E-04*
No Load	-6.674E-05 **	-6.232E-05 *	-5.918E-05 *	-6.702E-05 **	-6.651E-05 **	-5.967E-05 *	-5.317E-05
	3.291E-05	*3.531E-05*	*3.247E-05*	*3.287E-05*	*3.288E-05*	*3.265E-05*	*3.240E-05*

Table 6. Regressions of Alpha Components (Continued)

Factor	Base Model	TrueIndex	RelSize	FundsInIndex	Herfindahl	RelsizeFunds	RelsizeHerf
True Index		2.144E-05					
		6.128E-05					
Relative Size			1.105E-04 ***			1.041E-04 **	1.666E-04 ***
			4.090E-05			***4.720E-05***	***4.867E-05***
Relative Size*ETFDummy			-2.181E-04 ***			-2.174E-04 ***	-2.374E-04 **
			7.508E-05			***8.265E-05***	***9.689E-05***
Funds Following Index				-4.464E-07		-8.950E-08	
				2.890E-07		***3.282E-07***	
Funds Following Index*ETFDummy				1.166E-06		-3.431E-07	
				1.818E-06		***1.931E-06***	
Herfindahl Index					-8.876E-10		-8.647E-09 **
					3.518E-09		***4.121E-09***
Herfindahl Index*ETFDummy					-1.445E-08		3.495E-10
					1.019E-08		***1.319E-08***
Adjusted R-Squared	0.14797	0.14476	0.17687	0.14995	0.15026	0.17017	0.18683

Regressions were performed using the equation: $\hat{\alpha}_i = C_i + \sum_{j=1}^{n} \gamma_{i,j} V_{i,j} + \omega_i D_i + \sum_{j=1}^{n} \lambda_{i,j} V_{i,j} D_i + \eta_i$ to determine the significant components of abnormal return. Each factor (*V*) is listed in the left column. The "No Load" and "True Index" variables are not listed with interaction terms because they are only applicable to index funds. Standard errors are below the coefficients. *, **, *** represent the 10%, 5%, and 1% levels of significance, respectively.

Table 7. Spearman Correlations

PANEL A. All Pair-Types Excluding International Funds

n = 477

Minimum Absolute Correlation	Frequency	Percent	Cumulative Frequency	Cumulative Percent
99%	115	24.06%	115	24.11%
95%	313	65.48%	428	89.73%
90%	41	8.58%	469	98.32%
85%	5	1.05%	474	99.37%
<85%	3	0.63%	477	100.00%

PANEL B. International Funds

n = 30

Minimum Absolute Correlation	Frequency	Percent	Cumulative Frequency	Cumulative Percent
95%	0	0.00%	0	0.00%
90%	2	6.90%	2	6.67%
80%	1	3.45%	3	10.00%
70%	13	44.83%	16	53.33%
60%	7	24.14%	23	76.67%
50%	5	17.24%	28	93.33%
<50%	2	3.45%	30	100.00%

We perform a Spearman Correlation test on 507 combinations of indexes, index funds, and ETFs. Panel A combines all pair-type correlations except Indexes with international funds (both index funds and ETFs). The column "Frequency" shows the number of tests falling between the specified significance levels (i.e. for Panel A, 313 pairs have correlations between 95% and 99%).

Table 8. Tracking Error

	TE1	TE2	TE3		TE1	TE2	TE3
ALL ETFS (72)				***ALL INDEX FUNDS (177)***			
Mean	0.425%	0.600%	0.583%	Mean	0.261%	0.430%	0.212%
Std. Dev.	0.306%	0.406%	0.392%	Std. Dev.	0.682%	0.937%	0.198%
Median	0.257%	0.391%	0.379%	Median	0.057%	0.169%	0.163%
Skewness	0.851	0.764	0.801	Skewness	4.711	4.913	3.480
DOMESTIC ETFS (46)				***TRUE INDEX FUNDS (156)***			
Mean	0.216%	0.331%	0.324%	Mean	0.075%	0.179%	0.176%
Std. Dev.	0.092%	0.158%	0.152%	Std. Dev.	0.080%	0.123%	0.115%
Median	0.187%	0.285%	0.277%	Median	0.053%	0.151%	0.150%
Skewness	1.987	1.938	1.865	Skewness	5.781	2.942	2.472
INTERNATIONAL ETFS (26)				***OTHER INDEX FUNDS (21)***			
Mean	0.793%	1.078%	1.042%	Mean	1.887%	2.627%	0.533%
Std. Dev.	0.172%	0.228%	0.228%	Std. Dev.	1.475%	2.089%	0.443%
Median	0.750%	1.011%	0.994%	Median	1.586%	2.301%	0.377%
Skewness	1.261	1.147	1.340	Skewness	1.265	1.355	1.033

Descriptive statistics for each of the tracking error methods are presented. ETFs and index funds are listed separately. ETFs are separated into domestic and international funds. Index funds are separated into "true index funds" and other index funds. A fund is excluded from the true index fund category if it does not attempt to exactly mimic the index that it is following. For example, Rydex funds were excluded because they are attempting to achieve a beta of +/- 2.0 relative to their respective indexes.

Tracking error 1 (TE1) is calculated as $TE_{1,p} = \frac{\sum_{t=1}^{n} |e_{pt}|}{n}$, tracking error 2 (TE2) is calculated as $TE_{2,p} = \sqrt{\frac{1}{n-1}\sum_{t=1}^{n}\left(e_{pt} - \bar{e}_p\right)^2}$, and tracking error 3 (TE3) is the standard error of the regression that regresses the ETF or index fund on the index that the fund is attempting to follow.

Table 9: One- and Two-Period Fund Flows

Panel A: First period fund flows

	TE1	TE2	TE3
Psi (Constant)	1.274**	1.272**	0.585*
	0.570	***0.570***	***0.304***
Tracking Error	-27.198*	-21.028*	39.582
	16.007	***12.37***	***31.885***
ETF Dummy	0.398*	0.405*	0.401*
	0.226	***0.219***	***0.224***
ETF*Tracking Error	-35.090	-16.798	-11.811
	114.30	***79.345***	***83.047***
INTL Dummy	0.593*	0.606*	0.614*
	0.340	***0.331***	***0.332***
INTL*Tracking Error	-90.037	-72.060	-82.485
	105.30	***75.422***	***78.975***
True Index Dummy	-1.086*	-1.118*	-0.434
	0.572	***0.572***	***0.308***
TRUE*Tracking Error	147.370**	107.540***	51.914
	59.475	***40.631***	***52.195***
Adjusted R-squared	0.096	0.100	0.083
Number of observations:	691	691	713

Table 9: One- and Two-Period Fund Flows (continued)

Panel B: Second period fund flows

	TE1	TE2	TE3
Psi (Constant)	0.105	0.107	0.293
	0.231	***0.235***	***0.257***
Tracking Error	10.241	7.828	6.605
	9.148	***7.140***	***25.372***
ETF Dummy	0.101	0.076	0.119
	0.242	***0.208***	***0.189***
ETF*Tracking Error	125.370	119.580*	112.750*
	103.34	***66.998***	***60.897***
INTL Dummy	1.015***	1.010***	0.774**
	0.383	***0.357***	***0.326***
INTL*Tracking Error	-173.880*	-126.440*	-96.639
	103.68	***68.596***	***61.881***
True Index Dummy	0.003	0.034	-0.148
	0.232	***0.237***	***0.259***
TRUE*Tracking Error	23.025	-12.970	-14.526
	27.581	***16.020***	***29.351***
Adjusted R-squared	0.268	0.270	0.246
Number of observations:	449	449	472

One- and two-period fund flows were regressed on each type of tracking error with the regression:

$$FF_{t,p} = \psi_{t,p} + \kappa_{t,p} TE^{n}_{t-s,p} + \sum_{i=1}^{3} \tau_{i,t,p}\delta_i + \sum_{i=1}^{3} \theta_{i,t,p}\delta_i TE^{n}_{t-s,p} \ .$$

where FF_t is fund flow at time *t*; ψ, κ, τ and θ are parameters; *TE* is the tracking error variable at time *t–s*, where *s* = 1 and 2; *n* is the subscript denoting each method of calculating tracking error; δ_1 is a dummy that equals one if the fund is an ETF and zero if an index fund; δ_2 is a dummy that equals one if the fund is an international fund and zero if it is not an international fund; and δ_3 is a dummy that equals one if the fund is a true index fund and zero if it is not a true index fund. Standard errors are below the coefficients. *, **, *** represent the 10%, 5%, and 1% levels of significance, respectively.

Figure 1: Total Investment Company Assets

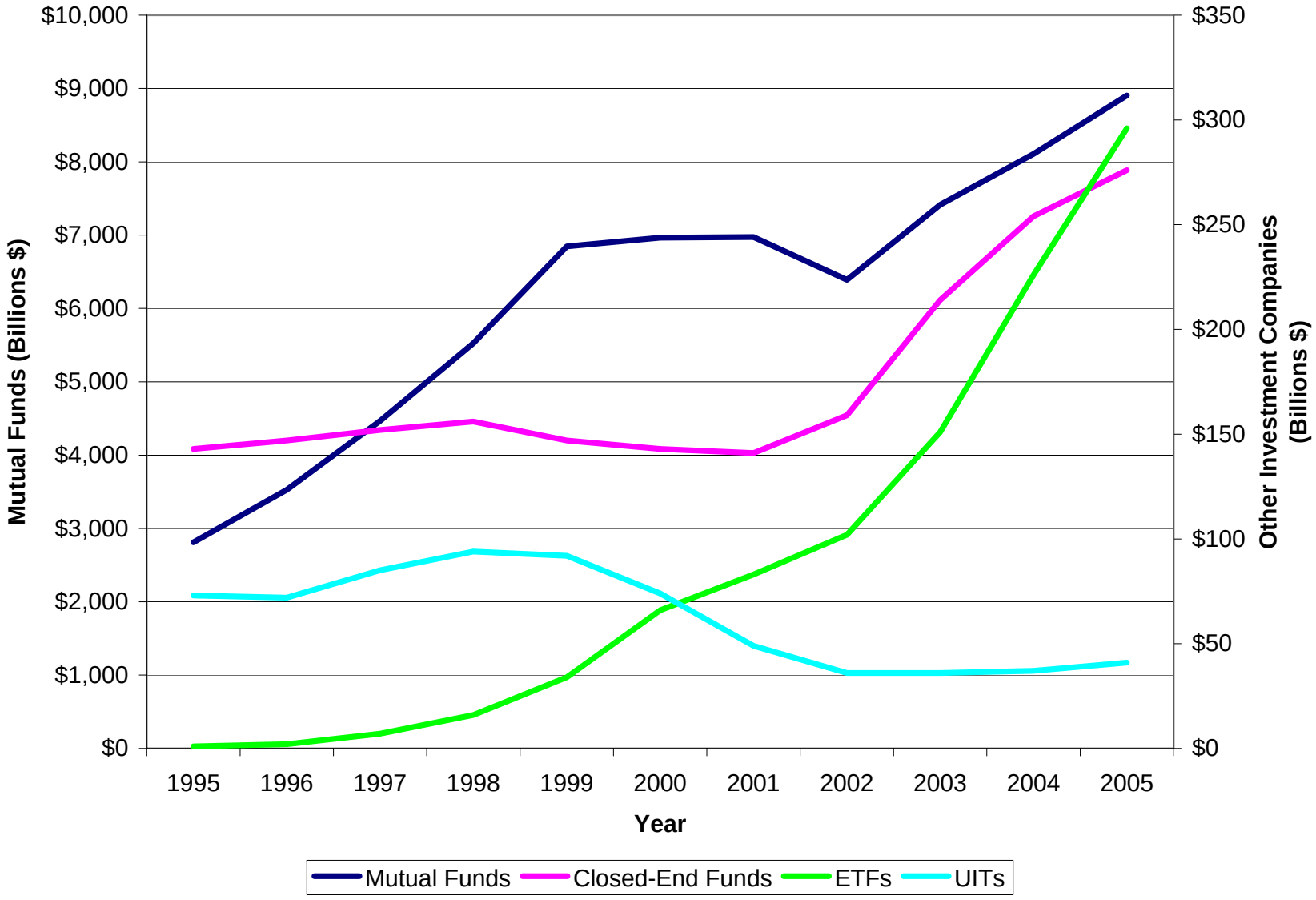

Figure 2: Net Fund Inflows