Index funds or ETFs: the case of the S&P 500 for individual investors

Larry J. Prather[a,*], Ting-Heng Chu[b], M. Imtiaz Mazumder[c], John C. Topuz[a]

[a]Department of Accounting & Finance, John Massey School of Business, Southeastern Oklahoma State University, Durant, OK 74701, USA
[b]Department of Economics and Finance, College of Business and Technology, East Tennessee State University, Johnson City, TN 37614, USA
[c]School of Business, SUNY Institute of Technology, Utica, NY 13504, USA

Abstract

We investigate alternative S&P 500 indexing strategies for individual investors using S&P 500 index funds and the Standard and Poor's depository receipt (SPDR). This investigation is important because while SPDRs have lower advertised annual expenses, investors in SPDRs face bid-ask spreads and commissions. We compute average spreads of SPDRs using transaction-by-transaction data, present a model to illustrate how investors can compare alternative index investments, and illustrate the results under several scenarios. We conclude by comparing risk-adjusted returns of the alternatives to ensure that undisclosed trading costs do not alter the choice. © 2009 Academy of Financial Services. All rights reserved.

1. Introduction

Empirical research on the performance of actively managed mutual funds suggests that they fail to consistently outperform the market on a risk-adjusted basis. Abundant studies find little evidence of market timing ability (e.g., Becker, Ferson, Myers, and Schill, 1999; Chang and Lewellen, 1984; Daniel, Grinblatt, Titman, and Wermers, 1997; Grinblatt and

---

* Corresponding author. Tel.: +1-580-745-2566; fax: +1-580-745-7485.
E-mail address: lprather@se.edu (L.J. Prather)

Titman, 1989, 1994; Henriksson, 1984; Kon, 1983; Lehmann and Modest, 1987; Treynor and Mazuy, 1966). Some studies even report evidence of perverse timing ability (e.g., Cumby and Glen, 1990; Viet and Cheney, 1982). Moreover, Gruber (1996) and Fortin and Michelson (1999, 2002) argue that the evidence is conclusive that actively managed funds under perform.

In 1976, the Vanguard Group launched the first S&P 500 index fund to provide a low-cost buy-and-hold alternative to active management. By mirroring the index, costs should be low because research expenses are eliminated, portfolio management requires fewer resources, and transactions costs (bid-ask spreads and commissions) are reduced. Moreover, given the emerging evidence at that time that portfolio managers incurred marginal costs in excess of the marginal returns they garnered, indexing appeared to be a better alternative to investors than investing in actively managed funds. Low cost indexing became increasingly popular with investors and by December 2006, 169 more S&P 500 index mutual funds were created by Vanguard's competitors. Thus, investors desiring to index all or parts of their portfolio to the S&P 500 have many mutual funds available from which to choose.

In 1993, State Street Bank & Trust Company introduced Standard & Poor's depository receipts (SPDRs) that track the S&P 500 index. The SPDR (ticker symbol SPY) is an Exchange Traded Fund (ETF) and is claimed to possess lower annual expenses than index funds. A competitor of the SPY is the iShares S&P 500 index (ticker symbol IVV) that was introduced in May 2000. However, the SPY is more popular, has more trading volume, and a longer return history. We concentrate on the SPY in this paper because the reported annual expenses of the SPY are lower than those of the IVV, thus, to the extent that these expenses represent relative costs of ownership, the SPY would be the superior investment. Unlike mutual funds, ETFs must be purchased through a broker; therefore, investors face brokerage fees to purchase and sell ETF shares. In addition, because trading ETFs is similar to trading shares of stock, ETFs are purchased at the dealer's ask price and sold at the dealer's bid price, which is lower than the ask price. Thus, ETFs resemble "load" mutual funds because load funds must be purchased at the public offering price [that is higher than net asset value (NAV)] and redeemed at NAV.

The dilemma facing investors who desire to implement an S&P 500 index strategy is whether index mutual funds or SPDRs are the better investment vehicles. Although SPDRs claim lower annual expenses, no-load index funds can be purchased and redeemed at NAV and have no trading costs. Thus, the answer depends largely on the magnitude of the bid-ask spreads of the SPDRs, differences in the magnitudes of expenses, commissions incurred in trading SPDRs, transaction size, and expected length of the investment.

To assist investors in determining the optimal investment, we investigate the reported annual expenses of competing S&P 500 index funds and the SPDR. We then compute average bid-ask spreads for SPDRs, using 94,442,386 bid-ask quotes during the 980-day trading period from February 7, 2000 through December 31, 2003. Next, we present a model to compare index investments with differing cost structures and illustrate results under several scenarios. Then we examine differences in portfolio management issues that can cause differing turnover, tracking errors, and unreported costs. We conclude by computing risk-adjusted returns of alternative index investments and model the results of the best index

mutual fund to the ETF alternative to ensure that the undisclosed trading costs do not alter the outcome.

The remainder of the paper is organized as follows. The following section discusses differential characteristics of mutual funds and ETFs. Section 3 discusses our data and methodology and Section 4 presents our results. Section 5 concludes.

2. Differential characteristics between ETFs and mutual funds

ETFs are designed to replicate the holdings, performance, and yield of their underlying index. ETFs are essentially index funds that are listed on an exchange, priced and traded intra day, and allow investors to buy or sell shares based on the collective performance of an entire portfolio. According to the Investment Company Institute, between 1996 and 2007, the number of ETFs increased from 19 to 629 and their net assets increased from $2.4 billion to $608.4 billion.

ETFs are created by large investors, and institutions, in block-sized units of shares known as "creation units." Creation units require a deposit with the trustee for a specified number of shares of a portfolio of securities that closely approximate the composition of the specific index. Similarly, block-sized units of ETF shares can be redeemed for a portfolio of securities approximating the index and a specified amount of cash. These creation units are one mechanism that helps to keep an ETF trading on the exchange at a price extremely close to the value of its underlying portfolio. Because ETFs are both created from the securities of an underlying portfolio and can be redeemed into the securities of an underlying portfolio, arbitrage opportunities would occur from price discrepancies between an ETF and the underlying portfolio. Thus, tracking errors between the ETF and the index should be small.

Market makers, broker-dealers, and specialists are the major players in the creation and redemption units of ETFs; however, many individual investors hold ETFs for risk-control reasons, diversification, and short-term trading. Elton, Gruber, Comer, and Li (2002) document that individual investors, broker dealers, and institutions with less than $100 million in assets are the predominant holders of SPDRs.

Among the advantages that ETFs claim are tax efficiency, lower annual costs, transparency, buying and selling flexibility, all day tracking and trading, and diversification. Several of these features may be especially important to individual investors. First, ETFs claim good tax benefits because they generate few realized capital gains. This should arise from the low turnover of the securities that comprise the portfolio. Low turnover should occur for two reasons. First, because an ETF is not actively managed, ETFs only sell securities to reflect changes in their underlying index. Moreover, when index composition changes require portfolio changes to be made, ETF investors are not taxed because the investor does not see a capital gain or loss. The ETF investor has a claim on the index value, not a particular stock. However, mutual fund investors are taxed on capital gains because the fund must sell the stock once it is dropped from the index. Exchange trading of ETFs further enhances their tax efficiency because investors who want to liquidate shares of an ETF sell them to other investors in the secondary market. Because of this structure, ETFs, unlike open-end investment companies (mutual funds), are not required to sell securities to meet redemptions. Thus, this structure eliminates the generation of trading-related capital gains that would be taxable

for remaining investors. ETFs also claim to have significantly lower annual expense ratios than mutual funds. According to Morningstar's Barclays Global Investors Analysis (May 2003) most iShares have expense ratios of approximately 60 basis points whereas actively managed funds and index funds average expenses of 192 and 106 basis points, respectively. However, the expense ratios of broad classes of ETFs are on average less than 20 basis points whereas, index mutual funds have expense ratios as high as 300 basis points.[1] This is because, in part, of the lack of a need to incur management expenses or trading costs to issue or redeem shares. Additionally, because ETFs are exchange traded, they can be bought and sold at intra day market prices, purchased on margin, sold short, and traded using stop and/or limit orders.

If the claims are true, the annual expenses of SPDRs should be lower than those of S&P 500 index funds. Moreover, because mutual fund trading costs are not included in published expense ratios, and those costs adversely affect fund performance, it is possible that the advantage of SPDRs is even greater than it appears. Despite the advantages, ETFs require investors to pay round trip commissions and a bid-ask spread. Therefore, the true cost of investing the SPDR cannot be compared directly to investing in an index fund and the question of which investment strategy is superior is an empirical issue.

3. Data and methodology

3.1. Data

We created an initial list of 170 competing S&P 500 index funds from the fund finder located on the World Wide Web at www.smartmoney.com in February 2007. Data on front-end load fees, contingent deferred sales charges, and annual expenses are from Lipper. Of the 170 funds, 14 funds are closed to new investors. Because our research focus is on the best S&P 500 index investment for individual investors, we examine the 156 funds that remain open to new investors. Of those 156 funds, 73 funds are institutional funds leaving 83 funds that individuals may find as suitable investments. Of the 83 funds, Morningstar classifies 23 as class "A" shares (front load), 16 as class "B" shares with decreasing back loads, 11 as level back loads (of 1%), and 33 as class "C" shares (no-load).[2] Assets are computed from Lipper data for the competing S&P 500 index funds as an illustration of the value of assets under management for each category. Daily return data from January 2, 2004 through December 31, 2007 for the funds, the S&P 500 index, and the ETFs are from the CRSP Survivorship Bias Free database and CRSP daily stock and index files, respectively. Table 1 provides descriptive statistics for our sample.

To determine the approximate bid-ask spreads that an investor is likely to encounter in trading SPDRs, we acquired historical bid-ask quotation transaction-by-transaction data from TickData. To compute bid-ask spreads, we use 94,442,386 quotes during the 980-trading day period from February 7, 2000 through December 31, 2003 and compute the spread as

$$Spread = \frac{Ask - Bid}{Ask} \qquad (1)$$

Table 1   Sample of S&P 500 index funds

Load type	Number of funds	Number open	Total/average assets (million $)
Back load	18	16	1,321.4/73.4
Back level	13	11	1,140.6/87.7
Front load	25	23	4,842.0/193.7
No load	37	33	175,365.6/4,739.6
Institutional	77	73	121,077.4/1,572.4
Total	170	156	303,747.0/1,786.7

Columns one through four present the types of load charges, the total number of funds in existence for each load structure, the number of funds open to new investors, and the assets invested in each category as reported by Lipper in February 2007.

3.2. Methodology

The choice between S&P 500 index funds and the SPDR can be framed in the same manner as the choice between class "A," class "B," and class "C" shares of a selected mutual fund. Because our sample includes a mix of class A, B, and C shares, we can evaluate all three classes of funds and the SPDR using a similar framework. The rationale for this common framework is that the investor is investing in the same underlying portfolio, S&P 500 index, and the only difference between the investments is the magnitude and type of costs and fees that must be paid to purchase, hold, or redeem the investment.

If one index fund had lower front-end load fees, back-end load fees, operating expenses, and 12-b1 fees, the choice would be easy because it would dominate all other index investment choices. Fortin and Michelson (1998) document that no load funds have, on average, higher percentage brokerage costs than load funds. However, lowering one fee generally requires trading off another. Thus, in the absence of dominance in fees, decision makers must examine which index alternative provides the highest net terminal wealth (NTW) by computing

$$NI \, (1 + NPR \,)^N = NTW, \tag{2}$$

where NI is the net investment, NPR is the net periodic return, and N is the number of investment periods. In the case of mutual funds that have a contingent deferred sales charge (back end load or class "B" shares), the terminal wealth must be multiplied by $(1 - BL)$ where BL is the percentage back-end load to compute NTW. For ETFs, the commission for selling the shares must be subtracted from the terminal wealth to compute NTW.[3] Risk differences can be ignored for this analysis because the investment is the same underlying portfolio.

Given any gross investment amount, the net investment and net periodic returns will differ between the investment choices if load and fee differentials exist. The net investment in class "C" shares is the investor's total (gross) investment because the investor does not pay a load. However, for class "A" shares, only part of the initial investment amount is invested and the remainder is the load fee. Thus, the net investment in a load fund is

$$NI = GI \, (1 - L), \tag{3}$$

where GI is the gross amount available for investment and L is the load percentage.

Modifying this approach to evaluate an ETF is straightforward. Because investors must purchase an ETF at the dealers' ask price and sell it at the dealers' bid price, this spread is similar to a front load. Additionally, purchasing an ETF involves a commission that is a flat fee. Thus, the NI for evaluating ETFs is

$$NI = (GI - C)(1 - S), \tag{4}$$

where GI is the gross amount to be invested, C is the commission, and S is the bid-ask spread percentage.[4]

To compute the expected net periodic returns, NPR, we need the expected gross returns on the underlying portfolio. Because we are using S&P 500 index alternatives, these should be the expected returns on the S&P 500 index. From the gross returns, operating expenses and 12b-1 fees should be subtracted to arrive at net returns. Total annual fees will differ among mutual funds, and between mutual funds and ETFs, causing net returns to differ even though the expected returns of the underlying portfolio will be identical.

If reported annual expenses were the only expenses an investor faced, our analysis would be complete. However, the annual expenses that are reported to investors do not include trading costs incurred by the fund. Fortin and Michelson (1998) show that trading costs are significantly and positively related to the percentage of brokerage costs. Moreover, these trading costs are not known a priori and will depend, in part, on the investment flows into and out of the fund. Because net returns can differ because of these undisclosed trading fees, we confirm our results by examining risk-adjusted daily excess returns using the single index model from January 2, 2004 through December 31, 2007. The single index model is computed as

$$R_p = \alpha + \beta(R_m) + \varepsilon, \tag{5}$$

where R_p is portfolio return, R_m is S&P 500 index return, α is the risk-adjusted excess return, β is the systematic risk, and ε is the error term.

4. Empirical results

4.1. Mutual fund expenses

Lipper data on mutual fund expenses as of February 2007 are summarized in Table 2. The expense ratio is expressed as the percentage of assets deducted each fiscal year for fund expenses. These include 12b-1 fees, management fees, administrative fees, operating costs, and all other asset-based costs incurred by the fund. However, portfolio transaction fees, or brokerage costs, as well as initial or deferred sales charges are not included in the expense ratio.

The data clearly show that there are significant differences in fees and that some funds dominate others. This is consistent with Haslem, Baker, and Smith (2006) who reveal a wide disparity in expense ratios. Thus, before attempting to compare all the potential investment choices, we can narrow the field if we can find the dominant fund for each load type. Table

Table 2   Fees and expenses of S&P 500 index funds

| | Load Type | | | |
	No load	Front	Back	Back level
Back load				
Maximum			5.00	
Minimum			3.00	
Average			4.56	
Back level				
Maximum				1.00
Minimum				1.00
Average				1.00
Front load				
Maximum		5.75		
Minimum		1.50		
Average		4.28		
Expenses				
Maximum	0.86	0.80	1.41	1.45
Minimum	0.07	0.25	0.91	1.00
Average	0.35	0.58	1.24	1.30

Columns one through five report the maximum, minimum, and average charge that investors would incur from investing in various classes of index funds. All charges are in percentage as reported by Lipper as of February 2007.

3 presents the load type dominance analysis. In the back load category, Munder Index 500 (B) (ticker symbol MUXBX) has the lowest back load and lowest annual fees; thus, it dominates all other back load funds. Rydex Index 500 (C) (RYSYX) exhibits similar dominance in the level load category. Interestingly, Munder Index 500 (B) commands only 5.7% of the funds invested in open funds in its category and Rydex Index 500 (C) commands only 0.6% of the funds invested in open funds in its category.

Dominance in front load funds is unclear because front load funds involve trading off lower front load fees for higher annual expenses. Thus, investors would need to evaluate the tradeoff between loads and fees, as shown in Section 3.2., to determine the optimal front-load

Table 3   Dominant funds by reported expenses

Category	Ticker symbol	Load	Expense	Minimum investment $	Inception date
Back load	MUXBX	3.00	0.91	2,500	10/31/95
Back level	RYSYX	1.00	1.00	1,000	5/31/06
Front load	PLSAX	1.50	0.66	1,000	6/28/05
	AEXAX	2.50	0.60	500	10/15/98
	GUSPX	3.00	0.53	5,000	8/27/00
	RYSOX	4.75	0.25	2,500	5/31/06
No load	UAIIX	0.00	0.12	1,000	3/28/00
	ETSPX	0.00	0.09	5,000	1/17/99
	FSMAX	0.00	0.07	100,000	10/14/05

Columns one through six present each type of load expense category, the ticker symbol of the fund with the lowest load and expenses in each category, the minimum investment in the fund, and the fund's inception date. All loads and expenses are in percentage as were reported by Lipper in February 2007.

fund. However, that analysis is unnecessary because no-load fund alternatives dominate all classes of load funds with lower annual fees. For example, Table 2 shows that no-load funds have the lowest average expense ratio (0.35%) and the lowest total expense ratio.

The optimal no-load fund is driven by the size of the investor's desired investment. For investments as small as $1,000, the investor can invest in United Assoc S&P 500 II (UAIIX) with annual expenses of 0.12%. However, with a $5,000 investment, the investor would appear to be better off in E*Trade S&P 500 Index (ETSPX) with annual expenses of only 0.09%. Finally, for investors with $100,000 to invest, Fidelity S&P 500 Index Advisor (FSMAX) offers annual expenses of only 0.07%.

4.2. SPDR spreads and expenses

Although it is not possible to predict the precise spread that an investor will face on any given trade, we computed the average bid-ask spread percentage for SPDRs using 94,442,386 quotations from February 7, 2000 through December 31, 2003. The average spread for our sample is 0.33265%. This spread will serve as our proxy for the actual spread that an investor would face for our ensuing analysis. Because of this spread, for SPDRs to dominate mutual funds, they need to have lower annual fees than the comparable funds.

The advertised expense ratio for the SPDR is 0.10% and therefore it dominates all but seven of the open S&P 500 index mutual funds. Of the seven no load funds with reported expenses lower than or equal to the SPDR, Lipper reports that only two funds have a minimum initial investment of $100,000 or less (ETSPX and FSMAX). Thus, if reported expenses were the only cost of ownership, both ETSPX and FSMAX would be better alternatives for individual investors than the SPDR (we will examine those costs later in the paper). However, FSMAX requires an initial investment of $100,000 making it an option to only those investors with sizable portfolios. The best choice for investors with $5,000 to invest would be ETSPX. For investors with between $1,000 and $4,999.99 to invest, UAIIX is the best mutual fund option. Although UAIIX has slightly higher reported expenses than the SPDR, investors do not need to pay trading costs or commissions to invest in the fund. Thus, the choice between UAIIX and the SPDR must be evaluated.

4.3. Analysis of competing index investments based on reported fees

To complete this analysis, we need to make an assumption concerning trading costs of the SPDR. Many online brokerage firms offer trades for $8 to $11 per trade; however, we base our analysis on a trading cost of $9.99 per trade that is offered by one of the large online brokers. In addition, we use the average bid-ask spread of 0.33265% computed previously.

Table 4 presents results of comparing investments in UAIIX and the SPDR (ticker symbol SPY). Because trading costs are fixed, trading costs have more influence on smaller and shorter-term investments. Therefore, our analysis examines two initial investment amounts. Panel A examines a $1,000 investment that represents the minimum initial investment required by UAIIX. Panel B examines a $4,999.99 investment. This breakpoint was chosen because it is just one cent smaller than the minimum required to invest in ETSPX (which dominates UAIIX and SPY) in terms of reported expenses (i.e., 0.10% for SPY and 0.09%

Table 4 Comparative analysis using advertised fees

Panel A: Initial investment $1,000

Return	5 Years		10 Years		20 Years		30 Years		40 Years	
	UAIIX	SPY	UAIIX	SPY	UAIIX	SPY	UAIIX	SPY	UAIIX	SPY
9%	1,530.17	1,501.23	2,341.43	2,304.56	5,482.29	5,419.33	12,836.40	12,725.76	30,055.54	29,864.69
10%	1,601.75	1,571.90	2,565.59	2,526.10	6,582.23	6,508.42	16,887.27	16,744.02	43,325.74	43,052.17
11%	1,675.97	1,645.19	2,808.87	2,766.55	7,889.77	7,803.03	22,161.37	21,975.42	62,248.47	61,855.71

Panel B: Initial investment $4,999.99

Return	5 Years		10 Years		20 Years		30 Years		40 Years	
	UAIIX	SPY	UAIIX	SPY	UAIIX	SPY	UAIIX	SPY	UAIIX	SPY
9%	7,650.85	7,607.16	11,707.13	11,656.29	27,411.41	27,355.99	64,181.89	64,183.33	150,277.38	150,570.48
10%	8,008.71	7,963.38	12,827.90	12,772.94	32,911.08	32,845.45	84,436.20	84,437.01	216,628.29	217,040.69
11%	8,379.83	8,332.80	14,044.34	13,984.88	39,448.77	39,370.80	110,806.61	110,805.34	311,241.73	311,818.15

The illustrations above are for the initial investment amounts listed in the panel headings and assume average returns on the S&P 500 index of 9%, 10%, and 11%. Illustrations use only the annual expenses, reported by Lipper in February 2007, that are disclosed to investors. Trading costs for the SPY are assumed to be $9.99 for each transaction (initial purchase and the sale at the end of the reported horizon). The bid-ask spread for the SPY is assumed to be 0.33265%.

for ETSPX). Thus, we examine both extremes of the investment spectrum where UAIIX clearly dominates all other index funds in terms of reported annual fees and compare the results to investing in the SPY.

Results in Table 4 Panel A suggest that with an investment of $1,000, UAIIX outperforms the SPY for 5, 10, 20, 30, and 40 year holding periods if the average return on the S&P 500 index is between 9% and 11%. The selection of returns for the S&P 500 is motivated by the historical average return of the S&P 500 index being approximately 10% per year over extended intervals. Table 4 Panel B reveals that for the SPY to outperform UAIIX, the investor must invest nearly $5,000 and hold the investment for more than 30 years.[5] The difference at the 30-year horizon is very small and thus extremely sensitive to the assumptions concerning trading costs and bid-ask spreads. However, investors could use the procedure in Section 3.2. to perform their own analysis with the assumptions that they desire.

One limitation of the preceding analysis is that mutual funds may incur more trading costs than the SPY, thus, an analysis using only reported fees that exclude these costs might be biased in favor of the mutual funds.

4.4. Turnover of S&P 500 index funds

If reported annual expenses were the only expenses that investors faced, our analysis would be complete. However, fund-trading costs are levied but not reported to investors a priori. If those trading costs are material, the costs could alter the outcomes of our analysis.

One source of trading costs is turnover. The turnover of index funds should be lower than the turnover of actively managed funds because index funds do not buy and sell stock in an attempt to beat the market. However, both index funds and ETFs experience some turnover because of changes in index composition. However, index fund managers also face a unique problem. To track the index, and its performance, the portfolio must be fully invested in the securities that make up the index. Unfortunately, if mutual fund outflows on a given day (redemptions) exceed inflows (investments), securities must be sold to meet the redemptions. Thus, mutual fund turnover is partially determined by the investor base and the structure of the mutual fund industry. Moreover, trading costs that arise from these trades are not included in reported annual expenses. Because ETF trading is similar to that of stocks, ETFs escape turnover created by portfolio management issues related to net new money inflows (outflows). Thus, ETFs are expected to have lower turnover and lower trading related costs.

The importance of these trading costs may be best illustrated with an example. Only one S&P 500 index fund (ETSPX) accessible to many investors in terms of minimum investment ($5,000) dominates the SPY in reported fees. ETSPX advertises annual expenses of 0.09% compared to annual expenses of 0.10% for the SPY. Thus, in the previous analysis, ETSPX dominated the SPY. However, ETSPX might have higher unreported trading costs than the SPY and those costs may alter the analysis.

Table 5 presents the turnover percentage of all classes of S&P 500 index funds. The differences among funds are large and average turnovers (by class) range from 6.8% to 12.3%. Moreover, the range of turnover in the no-load group is from 4% to 30%, which suggests that turnover differences might have an important effect on total costs. The reported turnover of UAIIX, ETSPX, and FSMAX are 10%, 4%, and 7%, respectively. ETSPX has

Table 5 Turnover of S&P 500 index funds by load type

Load type	Maximum turnover %	Minimum turnover %	Average turnover %
Back load	20	3	6.8
Back level	42	3	12.3
Front load	30	3	8.0
No load	30	4	8.0

Columns one through four present each load type and the maximum, minimum, and average turnover, in percentage, for the load type. Computations are from data reported by Lipper in February 2007.

the lowest advertised fees and the lowest turnover of any fund that accepts initial investments of $100,000 or less. As for SPY, the turnover is 4%. Thus, the turnover of the mutual funds equals or exceeds that of the SPY as suggested by the structure of the investments.

In addition to turnover, trading costs arise from growth in the mutual funds assets. These costs arise from uncertain new money inflows (outflows) and the indivisibility of the underlying assets that prevent being 100% invested. These market imperfections create portfolio management problems for mutual fund managers that make it nearly impossible to mirror the index. Moreover, the imperfections create trades that generate trading costs and these trading costs can vary with the fund's brokerage arrangements.

4.5. Tracking errors

If managers were able to mimic perfectly the S&P 500 index without incurring any costs, the R-squared would be 1, suggesting that the two series move perfectly with one another (zero tracking error). Thus, tracking error can be defined as $1 - R^2$.

Tracking errors of competing index alternatives are provided in Table 6. As expected, tracking errors of each index mutual fund is low (less than 4%). Surprisingly, the tracking error of the SPY is nearly 9% and more than twice as large as any of the mutual fund alternatives. Given the issues that mutual fund portfolio managers face, our a priori belief was that mutual funds should have higher tracking errors. This belief was motivated by the fact that the SPY did not need to hold cash for redemptions or need to trade stocks frequently. Thus, it would be expected to have lower turnover and lower cash holdings than a mutual fund alternative.[6]

Table 6 Tracking errors

	SPY	ETSPX	UAIIX	FSMAX
Pearson correlation	0.955	0.981	0.980	0.982
R^2	0.912	0.962	0.961	0.965
Tracking error	0.088	0.038	0.039	0.035

Column one presents the metrics of interest for the best S&P 500 index alternatives in columns two through five. Metrics were computed using daily returns for the 304 trading-day period of October 17, 2005 through December 29, 2006.

4.6. Risk-adjusted returns

To capture true costs, it is necessary to compare risk-adjusted returns. The advantage of using risk-adjusted returns is that if the index could be perfectly and costlessly replicated, a single index model regression would show an $R^2 = 1$, a $\beta = 1$, and a risk-adjusted excess return (α) of zero. To the extent that costs are incurred, the α will be negative by an amount equal to the percentage of costs. Thus, αs can be used to analyze relative performance.

Of the 83 index funds in our initial sample, seven funds were not listed in CRSP and 11 funds did not have return data for the entire period. Therefore, we analyzed risk-adjusted returns for 65 index funds. Table 7 repeats the analysis reported in Table 3 using daily risk-adjusted returns for the competing index investments during the 1,006 trading-day period of January 2, 2004 through December 31, 2007. We used the single index model to ascertain risk-adjusted returns (Eq. 5). Given that transaction costs exist, securities are not perfectly divisible, and that portfolio management complications arise that prevent perfect tracking of the index, we would expect R^2 to be high but less than 1. We also expect α to be slightly negative, reflecting the fact that investors' realized returns were below the index returns because of expenses incurred by the fund.

The average α, β, and R^2 for the 65 funds is -0.000030799, 1.000893, and 0.998, respectively. As expected, all mutual funds underperform the S&P 500 because of trading costs and annual expenses. Moreover, the negative αs of 43 of the 65 funds were statistically significant at the 5% level. Of the 65 funds, 20 had βs that were statistically greater than one and two had βs statistically less than one. Analysis of funds suggests that a tradeoff exists in front-load funds between load fees and annual expenses; however, the analysis produced clear dominance in back-load, back-level, and no-load categories. Moreover, Table 7 clearly reveals that no-load funds dominate all other funds categories. The top three no-load funds were HIDAX, UAIIX, and ETSPX.

Table 8 presents results for the best performing mutual fund and the ETF alternatives. Interestingly, unlike the mutual fund alternatives, both SPY and IVV have αs that are slightly positive (but not statistically significant). The bottom row presents the annualized under performance of each of the investments based on a 250 trading day year (computed as $(1+\alpha)^{250}-1$). One explanation for the positive αs for the ETFs is that the

Table 7   Dominant funds by risk-adjusted returns

Category	Ticker symbol	Load	α	Minimum investment	Inception date
Back load	MUXBX	3.00	$-3.9575E-05$	2,500	10/31/95
Back level	GRMCX	1.00	$-5.1866E-05$	2,000	10/23/03
Front load	KNIDX	5.00	$-2.1960E-05$	1,000	11/25/92
	GUSPX	3.00	$-2.4398E-05$	5,000	8/27/00
	AEXAX	2.50	$-2.4525E-05$	500	10/15/98
No load	HIDAX	0.00	$-4.9783E-06$	500	4/19/96

Columns one through six present each type of load expense category, the ticker symbol of the fund with the lowest load and highest risk-adjusted return in each category, the minimum investment in the fund, and the fund's inception date. All αs are in percentage and were computed using the single index model and daily data during the period from January 2, 2004 to December 31, 2007.

Table 8   Single Index Model Risk-adjusted returns

	SPY	IVV	HIDAX
α	0.000000007	0.000003905	−0.000004978
	(0.000)	(0.074)	(−0.341)
	[0.999]	[0.941]	[0.733]
β^*	0.9759	0.9746	1.0004
	(141.303)	(141.187)	(521.013)
	[0.000]	[0.000]	[0.000]
R^2	.9521	.9520	.9963
Tracking error	0.0479	0.0480	0.0037
Annual over/under performance	0.00000175	0.00097662	−0.00124373

The results of single index model using daily returns of the competing index investments for the 1,006 trading-day period of January 2, 2004 through December 31, 2007 are presented above. The index proxy is the total return of the S&P 500 index. t-statistics and p-values for the hypothesis test that the estimated parameter is different from zero are provided below the parameter estimates in parentheses and brackets, respectively. Annual under performance assumes 250 trading days per year and is computed as $(1+\alpha)^{250}-1$

* Statistical tests reveal that the SPY and IVV have βs less than one whereas the β of HIDAX is not statistically different from one. The t-statistics are −3.48, −3.37, and 0.190 for SPY, IVV, and HIDAX, respectively.

single index model does not use a risk-free asset. That may cause slight deviations in the security market line that may penalize higher β funds and reward lower β funds. Therefore, we repeated the analysis using the Jensen and Fama and French three factor (FF-3) models. Results of the Jensen models reveal no change in the dominant funds or the superiority of ETFs over mutual funds. The average α, β, and R^2 for the 65 funds using the Jensen model is −0.000030625, 1.0012, and 0.998, respectively. The number of funds with statistically negative αs, at the 5% level, decreased from 43 to 21 and 21 funds had βs that were statistically greater than one and two had βs statistically less than one.

As a final examination of risk-adjusted returns, we utilized the FF-3 model to ascertain whether differences in loading on the SMB and HML factors would affect our results. Results using the FF-3 model revealed that 45 funds had statistically negative αs and 32 funds had βs greater than one (all at the 5% level). The average α, β, and R^2 for the 65 funds is −0.000031714, 1.0024, and 0.998, respectively. The FF-3 model does little to improve results because only 32 fund had statistically significant SMB factors and 27 funds had statistically significant HML factors (at the 5% level). For about 25% of the funds, both the SMB and HML factors were statistically significant. As for dominance, there was no change in the overall result that the IVV and SPY dominated all mutual fund options. There was also no change in the dominant front load, back load, or back level load fund. The only change was that VFIAX moved ahead of HIDAX to become the best no-load fund. HIDAX moved to second place and UAIIX remained in third place. However, VFIAX requires an initial investment of at least $100,000, making it out of reach to many investors. Thus, we use HIDAX as the dominant fund for comparison to the SPDR. The only other significant change is that the αs of IVV and SPY became slightly negative, but not statistically significant.

4.7. Analysis of competing index investments based on annual under performance

Because turnover and trading related costs associated with asset base growth generate undisclosed costs to investors, we repeat the analysis reported in Table 4 using average annual under performance rather than reported costs. Table 9 provides illustrations that are for the initial investment amounts listed in the panel headings and assumes average returns on the S&P 500 Index of 9%, 10%, and 11%. Illustrations also assume that the total annual cost to investors is the average under performance using the single index model reported in Table 8. Trading costs for the SPY are assumed to be $9.99 for each transaction (initial purchase and the sale at the end of the reported horizon). The bid-ask spread for the SPY is assumed to be 0.33265%.

Results in Table 9 suggest that investors with only $1,000 to invest should select the HIDAX mutual fund if their expected holding period is less than 10 years and the SPY for holding periods 20 years or longer. For investors with $5,000 to invest, the HIDAX mutual or longer fund is the better option for an expected holding period of 5 years. For holding periods of 10 years or longer, the benefits of SPY start to become dramatic. Investors can use the procedure in Section 3.2. to perform analysis for other scenarios.

Our analysis suggests that the SPY is probably a more appropriate investment than an index mutual fund for most investors. This is because of the fact that we illustrated results using small investment amounts and the value of the SPY over the mutual fund grows with the size of the investment. In addition, the benefit of the SPY grows with time and investors with short investment horizons should probably not be in the stock market.[7]

5. Conclusion

We investigate alternative S&P 500 indexing strategies for individual investors by examining S&P 500 index funds and the SPDR ETF. We examine the characteristics of the open funds such as reported annual expense and load differences, and provide a model that can be used to ascertain the most appropriate indexing strategy. The choice between investing in index mutual funds and investing in SPDRs is also important because SPDRs have lower advertised annual expenses but investors in SPDRs face bid-ask spreads and commissions.

To determine the effect of bid-ask spreads, we examine spreads of SPDRs during a 980 trading-day period from February 7, 2000 through December 31, 2003. Using these computed spreads, and realistic commissions, we present a model that can be used to compare SPDRs and different classes of mutual funds.

Using our model and computed bid-ask spreads, we evaluate mutual funds and SPDRs using reported annual expenses. In addition, because the operational characteristics of mutual funds differ from those of ETFs, we also examine the turnover, tracking errors, and risk-adjusted returns of mutual funds and SPDRs. Finally, we model the investment decision using risk-adjusted returns.

Our results are important to individual investors in several ways. First, we show that a substantial amount of money is invested in load S&P 500 index funds although this seems to make no economic sense. Secondly, we provide a framework with which to compare

Table 9   Comparative analysis using realized as

Panel A: Initial investment $1,000

Return	5 Years HIDAX	SPY	10 Years HIDAX	SPY	20 Years HIDAX	SPY	30 Years HIDAX	SPY	40 Years HIDAX	SPY
9%	1,529.87	1,508.21	2,340.50	2,325.96	5,477.89	5,520.15	12,820.95	13,082.08	30,007.29	30,984.20
10%	1,601.43	1,579.14	2,564.56	2,549.34	6,576.99	6,628.36	16,867.12	17,208.45	43,256.83	44,660.91
11%	1,675.64	1,652.70	2,807.76	2,791.76	7,883.55	7,945.48	22,135.16	22,579.29	62,150.35	64,131.50

Panel B: Initial investment $4,999.99

Return	5 Years HIDAX	SPY	10 Years HIDAX	SPY	20 Years HIDAX	SPY	30 Years HIDAX	SPY	40 Years HIDAX	SPY
9%	7,649.31	7,642.82	11,702.42	11,764.05	27,389.40	27,863.89	64,104.60	65,978.67	150,036.15	156,301.99
10%	8,007.11	7,999.79	12,822.80	12,889.94	32,884.90	33,449.63	84,335.27	86,777.02	216,283.70	225,096.75
11%	8,378.18	8,370.55	14,038.80	14,111.81	39,417.67	40,088.38	110,675.58	113,847.96	310,751.11	323,285.76

The illustrations are for the initial investment amounts listed in the panel headings assuming average returns on the S&P 500 index of 9%, 10%, and 11%. Illustrations assume that the total annual costs to investors are the average under performance reported in Table 8. Trading costs for the SPY are assumed to be $9.99 for each transaction (initial purchase and the sale at the end of the reported horizon). The bid-ask spread for the SPY is assumed to be 0.33265%.

various classes of mutual funds and to compare index funds with ETFs. Third, we provide historical bid-ask spreads of SPDRs. Next, our risk-adjusted return results suggest that reported annual expenses are a poor proxy for the total cost of ownership. When only the reported annual expenses of the mutual funds are considered, mutual funds dominate the ETF alternative. However, when total annual costs are estimated from the single index model, the SPDR dominates mutual fund alternatives. Thus, using reported expenses to make investment decision is suboptimal.

Notes

1. Morningstar's Barclays Global Investors Analysis (May 2003) reported expense ratios as high as 300 basis points for index funds in general, we did not find any fund with expenses that high in our sample of S&P 500 funds.
2. From our perspective, it makes no economic sense to invest in load index funds and we were tempted to delete them from our analysis. However, a substantial amount of money is invested in load funds and no-load funds with total expenses greater than 0.25%. Analysis of our sample, based on Lipper data, reveals that only 78.7% of the $4,177,811.6 million that is invested in S&P 500 index funds is invested in true no-load funds (fees less than 0.25%). Front load funds hold $4,048.6 million, traditional back load funds hold $1,232.3 million, back level load funds hold $572.8 million, and no-load funds with fees higher than 0.25% hold $31,937.4 million.
3. We will deal with bid-ask spreads for ETFs later in the paper.
4. This framework is based on an extension of front load funds. In a front load fund, the investor pays a public offering price that is higher than the NAV that would be received if the fund was redeemed. The bid-ask spread for an ETF can be modeled in the same manner as a front load because the purchase price for the ETF is the dealer's ask price and the sale price will be the dealer's bid price. Whereas spreads will vary trade-to-trade, it is not possible to determine precisely what the spread will be on any given transaction a priori. It is also not possible to account for the possibility of a trade being executed "inside" the quoted spread (which would cause our spreads to overstate the true spread). However, investors can obtain a current spread before trading. Moreover, the historical spreads that we computed should provide a reasonable estimate of the actual spread an investor could expect on a given trade.
5. Throughout the remainder of our analysis, we ignore the potential differences in tax effects discussed in Section 2 because investors holding these investments in retirement accounts (tax sheltered or tax deferred accounts) will not see a differential current tax effect. Investors in taxable accounts will fare better with the SPY than our analysis indicates because of the tax advantage of ETFs over mutual funds. However, the benefit of the tax effect is complicated and depends upon the investor's Federal and State tax bracket; thus, it is beyond the scope of this research.
6. ETFs have two basic structures: *Management Investment Companies* and *Unit Investment Trusts*. A management investment company can immediately reinvest dividends and other distributions whereas unit investment trusts distribute dividends quarterly or annually. Because SPDRs are organized as unit investment trusts, track-

ing errors may occur from the dividend distribution process. Elton, Gruber, Comer, and Li (2002) report that cash dividend and management expenses are the primary reasons behind tracking errors in SPDRs. In addition, Gastineau (2004) reveals that ETF managers must wait until the market closes to know what creations or redemptions will occur before they can rebalance their portfolios the next day (at opening prices). However, mutual fund managers do not suffer from this problem. This could also induce a slightly higher tracking error for ETFs. Blume and Edelen (2004) made similar arguments and found that S&P500 index mutual funds had minimal tracking error.

7. Our risk-adjusted return results suggest the IVV ETF outperforms the SPY. The choice between the SPY and IVV would depend on the difference in spreads that an investor faced for a particular trade. Because the IVV has a lower trading volume, the theory of market making suggests that it has a larger bid-ask spread. However, data limitations preclude us from verifying IVV spreads. From a practical perspective, an investor can compute those spreads for a given trade for both the SPY and the IVV and use the procedure in Section 3.2. to decide between the competing ETFs.

References

Becker, C., Ferson, W., Myers, D. H., & Schill, M. J. (1999). Conditional market timing with benchmark investors. *Journal of Financial Economics, 52*, 119–148.

Blume, M. E., & Edelen, R. (2004). S&P500 indexers, tracking error, and liquidity, *Journal of Portfolio Management, 30*, 37–47.

Chang, E. C., & Lewellen, W. G. (1984). Market timing and mutual fund investment performance. *Journal of Business, 57*, 57–72.

Cumby, R., & Glen, J. (1990). Evaluating the performance of international mutual funds. *Journal of Finance, 45*, 497–521.

Daniel, K., Grinblatt, M., Titman, S., & Wermers, R. (1997). Measuring mutual fund performance with characteristic-based benchmarks. *Journal of Finance, 52*, 1035–1058.

Elton, E., Gruber, M., Comer, G., & Li, K. (2002). Spiders: Where are the bugs? *Journal of Business, 75*, 453–472.

Fortin, R., & Michelson, S. (1998). Mutual fund trading costs. *Journal of Investing, 7*, 66–70.

Fortin, R., & Michelson, S. (1999). Fund indexing versus active management: The results are... *Journal of Financial Planning, 12*, 74–81.

Fortin, R., & Michelson, S. (2002). Indexing versus active mutual fund management. *Journal of Financial Planning, 15*, 82–94.

Gastineau, G. (2004). The benchmark index ETF performance problem. *Journal of Portfolio Management, 30*, 96–103.

Grinblatt, M., & Titman, S. (1989). Portfolio performance evaluation: Old issues and new insights. *Review of Financial Studies, 2*, 396–422.

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

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

Haslem, J. A., Baker, H. K., & Smith, D. M. (2006). Are retail S&P 500 index funds a financial commodity? Insights for investors. *Financial Services Review, 15*, 99–116.

Henriksson, R. D. (1984). Market timing and mutual fund performance: An empirical investigation. *Journal of Business, 57*, 73–96.

Kon, S. J. (1983). The market-timing performance of mutual fund managers. *Journal of Business, 56*, 323–347.

Lehmann, B. N., & Modest, D. M. (1987). Mutual fund performance evaluation: A comparison of benchmarks and benchmark comparisons. *Journal of Finance, 42*, 233–65.

Treynor, J. L., & Mazuy, K. (1966). Can mutual funds outguess the market? *Harvard Business Review, 44*, 131–6.

Viet, E. T., & Cheney, J. M. (1982). Are mutual funds market timers? *Journal of Portfolio Management, 8*, 35–42.