

# S&P 500 Indexers, Tracking Error, and Liquidity

*A complex answer to profiting.*

Marshall E. Blume and Roger M. Edelen

**MARSHALL E. BLUME** is the Howard Butcher professor of financial management at the Wharton School of the University of Pennsylvania in Philadelphia, PA.  
[blume@wharton.upenn.edu](mailto:blume@wharton.upenn.edu)

**ROGER M. EDELEN** is managing director of research at ReFlow Management, LLC, in San Francisco, CA.  
[edelen@ReFlow.com](mailto:edelen@ReFlow.com)

**T**he pioneering studies of Harris and Gurel [1986] and Shleifer [1986] initiated an extensive literature that uniformly finds abnormal average returns when S&P adds a stock to its 500 index or drops a stock; most authors find a partial reversal in the days following the change. The first studies find average abnormal returns of 3% to 4%. Later authors find average abnormal returns of 8% to 10%. The latter include Beneish and Whaley [1996, 2002]; Blouin, Raedy, and Shackelford [2001]; Bos [2000]; Dhillon and Johnson [1991]; Edmister, Graham, and Pirie [1994]; Kaul, Mehrotra, and Morck [2000]; and Lynch and Mendenhall [1997].

Beneish and Whaley [1996] use intraday stock prices to show that the price of a typical stock does not adjust immediately when it is added to or dropped from the index, suggesting a profit opportunity. Beneish and Whaley end with the carefully worded conclusion that indexers could enhance their returns by trading “earlier in the announcement period (in the extreme, with market-on-open orders on the morning following the announcement)” [1996, p. 1929].

If an indexer had followed this strategy, prices should now adjust concurrently with or almost immediately after announcement. Yet Beneish and Whaley [2002] find that prices still do not adjust immediately when S&P changes its index, although the adjustment is faster than in their earlier study.

That there remains a delayed adjustment to changes in the S&P 500 indicates that the trading behavior of



indexers has not eliminated this lag in price adjustment. One possible explanation is that indexers, whose job it is to track the S&P 500, find that the expected enhancement in returns is not sufficient to compensate for potential tracking errors. (Tracking error is defined as the absolute difference between the return realized by an indexer and return on the index—absolute in that both positive and negative errors indicate a deviation from the index return.)

We cannot assess this explanation from the analysis of Beneish and Whaley, as they do not translate the abnormal returns they find into variables of direct interest to indexers—the expected enhancement in returns and the potential tracking errors.

We first show that an indexer that follows an early trading strategy of trading at the opening price on the day following any announcement of change would have enhanced its return from 1995 through 2000 by an average of 19.2 basis points per year. At a standard deviation of 23.9 basis points per year, this tracking error is much greater than that of the largest indexers: for Barclays Global Investors, the largest indexer, just 2.8 basis points, and for Vanguard, the largest indexer of mutual funds, just 11.7 basis points per year. Yet even with these lower tracking errors, both firms have enhanced their returns—suggesting a more active strategy than just replicating the index.

We argue that an indexer that wants to maintain tracking errors in the same range as Barclays or Vanguard must very closely follow a full or exact replication strategy. An exact replication strategy requires holding every stock in the index with portfolio weights proportional to the weights in the index at every time. Market trading mechanisms now allow an indexer to trade at the closing price on a change day and be “paid” for entering into such a trade. Some indexers have thereby been able to share some of the profits that liquidity providers expect to earn, while minimizing tracking error.

## ENHANCING AN INDEX RETURN

We address three questions concerning the early trading strategy of trading at the open following the announcement of a change in the S&P 500 index rather than at the close on the day of the change. First, what is the average enhancement in return? Second, what is the tracking error? Third, by how much does this early trading strategy change the variance of total return and the covariances with other asset classes? This last question is important in constructing efficient portfolios.

## S&P 500

The S&P 500 is widely publicized, broad-based, and frequently used to benchmark returns. The market value of the 500 stocks in the index was \$11.7 trillion as of the end of 2000, or 68.7% of the market value of U.S. publicly traded equity.<sup>1</sup> The S&P 500 annual *Directory* includes a list of the leading S&P 500 fund managers and their assets under management. These assets totaled nearly \$1 trillion as of the end of 2000 (\$870 billion in non-enhanced products and \$63 billion in enhanced products) and represent 8.0% of the total market value of the index. The management of index funds is highly concentrated; the top three indexers manage 58% of all S&P 500 indexed assets.

Each stock in the S&P 500 is weighted in proportion to the market value of its common stock. Unlike other broad indexes, such as the New York Stock Exchange Composite, the Wilshire 5000, or the Russell 2000, the criteria for inclusion are subjective.

S&P regularly drops and adds companies. From 1995 through 2000, S&P announced 235 changes to the components of its index. S&P also periodically changes the number of shares outstanding of a component company that it uses in calculating the index. For large changes, such as 5% or more, S&P updates the number of shares immediately. For small changes, it waits until the end of the calendar quarter to update the index.

To calculate the value of the index, S&P prices each stock by the price of the last transaction on the primary market, not the consolidated closing price reported in the newspapers or by the Center for Research in Security Prices. For Nasdaq-listed stocks, the two prices are identical; for NYSE-listed stocks, the two prices can differ if the last trade occurs on a market other than the NYSE.

## Data

The primary data sources we use to calculate the returns of the early trading strategy are Vestek, CRSP, and the S&P *Directory*. The Vestek database provides for each stock in the index the month-end data: CUSIP number, shares outstanding, and month-end closing price.<sup>2</sup> The composite closing prices come from CRSP. The S&P *Directory* is the primary source for the announcement and change dates.

The sample covers the six years 1995 through 2000. The securities in Vestek and CRSP are first matched by CUSIP numbers, and then each match is checked by comparing the closing price on Vestek with the closing price



on CRSP. When a match fails this second test, we search the closing prices from CRSP to find the correct company.

### Early Trading Strategy

From 1995 through 2000, the price returns (the return without dividends) for the early trading strategy exceed the corresponding returns on the S&P 500 by an average of 1.6 basis points per month.<sup>3</sup> The standard deviation of the monthly tracking error is 6.90 basis points, implying a t-statistic of 1.97. On an annualized basis, the additional return is 19.2 basis points, and the standard deviation of the tracking error is 23.9.

Even though the annual tracking error can be high, the early trading strategy would have enhanced returns in five of the six years from 1995 through 2000. In 1995, the return on the early trading strategy of 34.24% exceeded the return on the index of 34.11% by 13 basis points. The annual tracking error was thus 13 basis points. The annual tracking errors for the subsequent years were: 28.9 bp in 1996, -5.4 bp in 1997, 43.9 bp in 1998, 12.9 bp in 1999, and 32.1 bp in 2000.

The correlation between the monthly returns of the early trading strategy and the S&P 500 is 0.9998683. The standard deviations of monthly returns of the two series are 4.2541% for the early trading strategy and 4.2528% for the S&P 500—the same to three significant figures. That portfolios with such similar total returns could result in a standard deviation of tracking error of 6.90 basis points per month needs explanation.

The size of the tracking error is driven almost exclusively by the less-than-perfect correlation between the returns on the early trading strategy and those on the S&P 500 itself, and not by the differences in standard deviations of the two return series. That is, even if the standard deviation of the returns for the early trading strategy were the same to four decimal places as the standard deviation of total return of the S&P, namely, 4.2541% per month, the usual formula for the standard deviation of the difference of two random variables with a correlation of 0.9998683 shows that the standard deviation for the tracking error is still 6.90 basis points to two decimal places.

The overall risk characteristics of the returns for the early trading strategy and the returns on the index are virtually the same. The standard deviation of either return series is the same to three significant figures—4.25% per month, and the covariances of either return series and the returns of any other asset are virtually the same.<sup>4</sup>

In mean-variance space, the early trading strategy

with its enhanced return virtually dominates the index. Yet, index fund managers may avoid an early trading strategy, as tracking errors with a standard deviation of 23.9 basis points per year may be unacceptable.<sup>5</sup>

### INVESTMENT STRATEGIES

An exact replication strategy requires holding all 500 stocks at all times in exact proportion to their weights in the index. Since changes to the index take place at the closing prices of the primary market, an indexer that wants no tracking error must buy the added stock and sell the eliminated stocks at these closing prices. The indexer would also have to adjust its holdings in each of the other 499 stocks, except when the market values of the added and deleted stocks are identical, an unlikely event. Similarly, a change to the index such as a reduction in shares outstanding when a company repurchases its own stock would also require trading in all 500 stocks.

Sampling strategies include maintaining portfolios of fewer than 500 stocks. These include stratified sampling, holding only the largest stocks, and optimization procedures. These strategies are likely to produce greater tracking error, but have the potential to reduce trading costs, broadly defined, leading to enhanced returns.

#### Actual Tracking Errors

The choice between exact replication and sampling is driven by the trade-off an indexer faces between enhancing returns and minimizing tracking errors. In practice, the largest indexers track the index quite closely. Barclays Global Investors is the largest manager of S&P 500 index funds. As of the end of 2000, it managed \$212 billion in S&P index funds, or roughly one-quarter of all money invested in S&P 500 index funds.

Barclays reports that from 1991 through 2000, the tracking error of its Equity Index Fund before management expenses averaged 2.2 basis points per year (*Exhibit 1*). The annual tracking errors are always close to zero, with a maximum error of 7 basis points in 1997 and a minimum error of -3 basis points in 1990. The standard deviation of this tracking error is 2.8 basis points per year. The positive average tracking error of 2.2 basis points means that Barclays enhanced its return over these years—a possibly surprising result, particularly since the Barclays returns are after trading costs. A spokesman for Barclays stated that it enhances returns by lending securities and by “smart” trading.<sup>6</sup>

The tracking errors for Vanguard, the third-largest



## EXHIBIT 1

### Tracking Errors of Barclays and Vanguard Against the S&P 500—1990-2000

Year	Barclays Before Expenses	Tracking Errors (percent per year) Vanguard 500		
		Before Expenses	Expenses	After Expenses
2000	0.00	0.23	0.18	0.05
1999	0.03	0.21	0.18	0.03
1998	0.04	0.22	0.18	0.04
1997	0.07	0.02	0.19	-0.17
1996	0.03	0.01	0.20	-0.19
1995	0.05	0.22	0.20	0.02
1994	0.01	0.06	0.19	-0.13
1993	0.04	0.09	0.19	-0.10
1992	-0.01	-0.06	0.19	-0.25
1991	0.01	-0.13	0.20	-0.33
1990	-0.03	0.07	0.22	-0.15
Simple Average	0.022	0.085		-0.107
Average Absolute Deviation	0.029	0.120		0.133
Standard Deviation	0.028	0.117		0.122
Assets under Management December 2000 (billions of dollars)	212			148

S&P 500 indexer after State Street Global Advisors, and the largest indexer of mutual funds, are somewhat higher than the Barclays tracking error, but still small. From 1991 through 2000, the standard deviation of tracking error of its Vanguard 500 before expenses was 11.7 basis points per year. The average tracking error is 8.5 basis points per year, and is positive in all but two years. In the last three years, Vanguard enhanced its return sufficiently to cover all its expenses.

The higher Vanguard tracking errors suggest that Vanguard deviates more than Barclays from an exact replication strategy, and through this deviation has successfully enhanced its returns. Consistent with this observation, Vanguard states in its prospectus that it will use derivatives when “favorably priced,” and it presumably uses “smart” trading techniques as well.

These tracking errors are low in comparison to those of more active management styles. Of the equity mutual funds operating throughout 2000, Morningstar classifies 922 non-index funds as large blend, the same classification it assigns to S&P 500 index funds. Of these 922 funds, 76 have betas with respect to the S&P 500 of from 0.98 to 1.02 and an R-square of 90% or more, suggesting that they are benchmarked to the S&P 500. The average tracking

error of these 76 funds is -34 basis points, which indicates lower returns than the index. The standard deviation of their annual tracking error is 383 basis points.

### Strategies of Mutual Funds

For the most part, mutual funds indexed to the S&P 500 hold every stock in the index, which is consistent with an exact replication strategy.<sup>7</sup> As of December 31, 2000, the Morningstar database included information on 2,992 domestic diversified equity funds with a three-year track record and equity style indicator. Their combined assets total \$2.5 trillion.

Of these 2,992 funds, 82 have in their name the words *index* or *S&P 500*, or some variant. Morningstar reports that all have a beta with respect to the S&P 500 of between 0.98 and 1.01, and all have an R-square of 1.00. The assets of these funds represent 9.3% of the assets of these 2,992 domestic diversified equity funds.

Of these 82 funds, 63 hold between 499 and 502 stocks; 14 hold more than 502 stocks, but none holds more than 507 stocks. Three funds hold 466 to 487 stocks, and two (two classes of shares for the same fund) hold only 240 stocks. These numbers are close to what one would expect if the investment strategies of these S&P index funds approximate an exact replication strategy.<sup>8</sup>

### THE LOGIC OF EXACT REPLICATION

That most mutual funds indexed to the S&P 500 hold approximately 500 stocks is no accident. It is virtually impossible to maintain tracking errors in the neighborhood of the 2 to 3 basis points as reported by Barclays without holding all 500 stocks in close proportion to the weights in the S&P 500.

Even if a portfolio has a seemingly high R-square with respect to the S&P 500, tracking errors can be high. For example, consider a portfolio with an R-square of 0.99 with respect to the S&P 500 and a beta of 1.0. If the standard deviation of the S&P 500 is 20% per year, which is close to its historical value from 1926 through 2000, routine calculations show that the standard deviation of the portfolio return is 20.1% per year, virtually identical to 20%. But the standard deviation of the tracking error is 201 basis points per year, implying much greater tracking errors than those of Barclays or of Vanguard. Even with an R-square of 0.9999, the standard deviation of the tracking error is still 20 basis points per year.

We describe in the appendix a model that demon-



strates the extreme sensitivity of tracking error to slight deviations from index weights. The model uses the market weights of each of the 500 stocks in the S&P 500 as of the end of 2000, and assumes a plausible single factor return-generating function for each stock. In this model, the standard deviation of annual returns for the single factor is 20% per year, and the standard deviation of the S&P 500 is 20.28%.

One sampling strategy is to invest in the largest 100 stocks. According to the model, such a strategy has a standard deviation of 20.53%, which is not much higher than the 20.28% for the index. Yet the standard deviation of the tracking error is 145 basis points, even though the R-square of this strategy with respect to the index is 0.9951.

Even dropping one stock from the index may result in significant tracking error. Dropping Nabor, the largest of the smallest 250 stocks, leads to a tracking error with a standard deviation of 2.6 basis points per year. The standard deviation of this portfolio is the same as the index to four significant figures.

An indexer that wishes to maintain tracking errors as low as those of Barclays or Vanguard must basically invest in ways that closely approximate an exact replication strategy. Otherwise, the tracking errors are likely to be higher than those observed in practice.

## LIQUIDITY AT THE CLOSE

Barclays and Vanguard have enhanced returns while maintaining a low standard deviation of tracking errors. To maintain such low tracking errors requires that their strategies have to quite closely replicate an exact replication strategy. The question then is how they can enhance returns while maintaining such low tracking errors. Trading in derivatives and security lending may be part of the answer.

Interviews with some practitioners have suggested an additional way. Counterparties such as hedge funds or dealers can enter into bilateral agreements with indexers to trade at a yet unknown closing price on the change date and agree to share part of their expected trading profits with the indexers through a better net price than the closing price. These counterparties are effectively paying indexers for the commitment to trade at the closing price. An indexer can thus enhance its returns with some certainty that it will not introduce negative tracking error.

## Trading Procedures

The closing mechanisms on Nasdaq and the NYSE differ. The NYSE has a formal mechanism that determines the closing price and guarantees an investor an execution at that price. Nasdaq has no such mechanism, but defines the closing price as the price of the last reported trade.<sup>9</sup>

There is no direct way for an investor to participate in this trade. An investor who wishes to trade at the closing price must enter into a bilateral agreement with a counterparty. After observing the closing price, the counterparty trades the agreed number of shares at that price.

Given the differences in these closing mechanisms, we analyze the two markets separately.

## Nasdaq

On Nasdaq, a counterparty that enters into a bilateral agreement to trade with an indexer at the closing price can obtain the required shares in several ways: buy on its own account the agreed-upon shares prior to the close on the change date; sell shares short at the closing price on the change date and cover later; or arrange for another investor like a hedge fund to provide the shares. Similarly, in the case of a stock deletion, the counterparty can sell short prior to the close and cover by buying from the indexer; buy from the indexer at the close and sell later; or arrange for a third party to buy the indexer's shares.

According to practitioners we interviewed, counterparties sometimes compensate indexers for entering into bilateral agreements to trade at the close, and our empirical evidence confirms this observation. If a counterparty is willing to pay, it must be the case that it receives some type of benefit. One possible benefit is that a counterparty that intends to take a speculative position would know to whom that position can be unwound at the closing price. This is a real benefit, as Nasdaq has no formal mechanism to assure that a non-prearranged trade would take place at the closing price.

Another possible benefit is that a counterparty that enters into such an agreement to trade a large position knows the trading strategy of that indexer. With that private knowledge, the counterparty may be able to extract more information from publicly reported volume and prices.

For example, if there is a print of a large trade a couple of hours prior to the close, the counterparty may be better able to assess the future trading strategies of other traders and indexers. A large indexer would require compensation to part with this private knowledge.



## NYSE

Unlike Nasdaq, the NYSE has a formal closing mechanism or auction that guarantees the execution of a previously submitted market-on-close (MOC) order. Thus, indexers do not need to enter into bilateral agreements with counterparties to trade NYSE stocks at the closing price, as they do for Nasdaq stocks. At the same time, practitioners indicate that such bilateral agreements exist and that indexers are paid to enter into these agreements. Our evidence confirms such bilateral agreements.

On the NYSE, agents can provide liquidity to indexers by entering MOC orders rather than using bilateral agreements, and be assured of execution at the NYSE closing price. The question then is what would induce a counterparty to pay an indexer to enter into an agreement to trade at the closing price, rather than participate in the NYSE closing auction.

Consider an addition to the index and an agent that seeks to profit from providing liquidity to large indexers. In submitting an MOC sell order, that agent faces uncertainty in that the MOC price will depend on the trading strategy of large indexers. The agent might expect a closing price of  $P_h$  if a particular large indexer submits an MOC buy order, but a lower price  $P_l$  if the indexer follows some alternative trading strategy. By entering into a bilateral agreement with the indexer, the agent has the assurance that the indexer will buy a specific number of shares at the close, resulting in an expected closing price of  $P_h$ . Again, as on Nasdaq, a counterparty that enters into a bilateral agreement with a large indexer obtains potentially valuable knowledge of the trading strategy of that indexer.

In principle, the two parties could submit their matched orders to the floor of the NYSE for execution at the close, but our evidence suggests that at least some of these bilateral agreements are executed and reported on non-NYSE markets. One possible reason is that one or both may not want to expose their large orders to the floor of the NYSE if such an exposure might have an adverse effect on the closing price. Another possible reason is that it is easier to effect a bilateral agreement on a non-NYSE market, as there is some anonymity on the NYSE floor.

### Some Institutional Detail

To establish the operation of bilateral agreements and payments for entering into such agreements, we use specific institutional details in the execution of trades and the reporting of their prices. If a dealer trades from its own

inventory with a public investor, a so-called *principal trade*, the dealer charges no commission, and profits solely from the spread. The reported price is the net price. If a dealer facilitates a trade between two public investors, a so-called *agency trade*, the dealer would normally charge a commission. The price reported on the tape is the gross price, before commissions. The net or actual price to either party is the gross price adjusted for any commissions.

During the years of this study, most trades in Nasdaq stocks were principal trades, a fact we can use to establish that payments are made to enter into bilateral agreements. To illustrate in the case of an addition to the index, an indexer might enter into a bilateral agreement with a dealer to buy a predetermined number of shares at the closing price less a “payment” of 5 cents. If the closing price turns out to be \$20, the indexer would buy the shares at the better price of \$19.95. This trade would be reported after the close with a condition code of T on the TAQ database at \$19.95. Thus, a substantial volume of trading in the case of additions that occur at a lower price than the closing price indicates such payments to indexers. The reverse applies to deletions. In short, indexers obtain a better price than the closing price.

Evidence of bilateral agreements for NYSE stocks can be established by looking at the specific reporting procedures for trades in these stocks. The NYSE’s formal mechanism to determine the price and volume of the last trade determines the closing price. After the report of this trade, no further trades in NYSE stocks can be reported on the NYSE; they must be reported through some other venue, primarily regional exchanges or Nasdaq. Thus, trades subject to bilateral agreements would be reported on these other venues after the close on the NYSE. If most trades of NYSE stocks are agency trades, the price of most of the trades reported after the close on the NYSE would be at the NYSE closing price.

### Evidence

For Nasdaq stocks, an exact replication indexer that entered into an agreement to buy or sell at the close but at a better price than the closing price would effectively be receiving a payment. When an indexer is buying at the close, a better price from the indexer’s point of view is a price lower than the closing price.

Consistent with the presence of such payments, 61% of the closing volume for additions of Nasdaq stocks is reported at a price lower than the closing price (see Exhibit 2).<sup>10</sup> Approximately 30% is reported at the clos-



## EXHIBIT 2

### Closing Volume as Percent of Outstanding on Change Day for Changes to the S&P 500—1996-2000

	Additions			Deletions		
	NYSE	Nasdaq	Total	NYSE	Nasdaq	Total
<b>A. Average volume at the close</b>						
40 days prior to announcement	0.01	0.01	0.01	0.01	0.01	0.01
Observations	124	37	161	41	5	46
<b>B. Announcement preceding change by 2 or more trading days</b>						
Actual volume at close on change day	3.58	5.84	4.05	4.43	6.47	4.65
Source of closing volume						
NYSE	2.46	0.00	1.95	3.52	0.00	3.14
Non-NYSE markets	1.12	5.84	2.10	0.91	6.47	1.51
Nasdaq	0.93	5.84	1.95	0.78	6.47	1.40
Regional exchanges	0.19	0.00	0.15	0.13	0.00	0.11
Percentage breakdown of non-NYSE closing volume						
Percent at better price	5.0	60.7	16.6	3.3	70.0	10.6
Percent at close price	94.3	29.4	80.8	95.2	28.0	87.9
Percent at worse price	0.7	9.9	2.6	1.4	2.0	1.5
Observations	106	28	134	41	5	46
<b>C. Announcement with change the following day</b>						
Actual volume at close on change day	4.33	5.66	4.77			
Source of closing volume						
NYSE	2.95	0.00	1.97			
Non-NYSE markets	1.39	5.66	2.81			
Nasdaq	1.36	5.66	2.79			
Regional exchanges	0.02	0.00	0.01			
Percentage breakdown of non-NYSE closing volume						
Percent at better price	7.0	58.9	24.3			
Percent at close price	92.5	29.5	71.5			
Percent at worse price	0.4	11.6	4.1			
Observations	18	9	27			

ing price, which is consistent with the waiving of commissions. Only 9.9% of the closing volume is reported at a price higher than the closing price, which is consistent with the payment of commissions.

When an indexer is selling at the close, a better price from the indexer's point of view is a price that is higher than the closing price. Again consistent with the presence of payments, 70% of the closing volume is reported at a price higher than the closing price. These results are consistent with compensation from counterparties to indexers to participate in trades at the closing price.

Evidence of bilateral agreements to trade NYSE stocks at the NYSE closing price would be established if there were a substantial volume in NYSE stocks reported after the NYSE close on other markets at the NYSE closing price. Consistent with this observation, roughly one-third of the closing volume for additions of NYSE stocks is reported on non-NYSE markets, and roughly one-quarter for dele-

tions. Moreover, approximately 95% of the volume reported on non-NYSE markets is at the NYSE closing price.<sup>11</sup>

Of those few trades that are reported with prices different from the NYSE close, most are at better prices from the indexers' perspective. This finding is consistent with a conclusion that dealers are the counterparty and pay indexers to enter into such trades. We have no data that would allow us to determine what payments, if any, are made to indexers for agreeing to trade at the NYSE closing price after the NYSE close. The evidence all the same is consistent with the operation of bilateral agreements to trade NYSE stocks at the closing price.

### Postscript

Our evidence suggests that many indexers closely follow an exact replication strategy. We have argued that the low tracking errors of Barclays and Vanguard require



a close approximation of an exact replication strategy. That virtually all mutual funds that are indexed to the S&P 500 hold approximately 500 stocks is also evidence consistent with following an exact replication strategy.

The volume at the close is also consistent with the proposition that many index funds, at least in terms of dollars under management, closely follow exact replication strategies. The dollar value of assets indexed to the S&P 500 is 8% of the total market value of the index, and the volume at the close for any added or deleted NYSE-listed stock is roughly 4% of that stock's market value. If index funds were involved in all these trades, these numbers indicate that half of all S&P 500 indexed assets are managed to approximate an exact replication strategy. Of course, if some of these closing trades involved non-indexers on both sides, this would moderate our conclusion.<sup>12</sup>

## CONCLUSION

On average, stocks that S&P adds to the S&P 500 index experience abnormal positive returns from the time of the announcement to the close on the change date, and then a partial reversal after the change date. The reverse occurs for stocks that are dropped. The adjustment following the announcement is not immediate, suggesting a profitable trading opportunity.

We show that an indexer who tried to take advantage of these abnormal returns by adjusting its portfolio to a change in the index immediately at the opening price on the day following the announcement of a change, rather than waiting until the close on the change day, would have enhanced its return by an average of 19.2 basis points a year. But the cost is that the tracking errors of such an early trading strategy are much higher than those observed in practice. Thus, an indexer who wants low tracking error would not pursue such an early trading strategy.

So who should profit from these expected price movements? It turns out that the answer is complex, and involves an interplay between indexers who demand substantial liquidity at the close on a change day and the suppliers of that liquidity. For example, counterparties that provide liquidity at the close for a stock addition can either buy the shares before the change and sell them to indexers at the close, or sell the shares short to the indexers and cover at a later date. To provide these services, counterparties must perceive that the expected returns from their long or short positions are sufficient to compensate for the risk to which they are exposed.

Studies of abnormal returns of changes to the index

have focused on averages, but there is a great deal of dispersion around these averages. Wurgler and Zhuravskaya [2002] show that there are only imperfect ways to hedge the risk of taking a position in a stock that will be added to or dropped from the index. If so, any counterparties providing liquidity will be facing significant risks.

As the liquidity needs of indexers are limited, essentially a scarce resource, it is possible that their counterparties will compensate indexers to enter into bilateral agreements to trade at the close. Of course, entering into a bilateral agreement with a large indexer discloses to the counterparty the trading intentions of that indexer—potentially valuable information.

As long as there are significant dollars invested to match the S&P 500 with minimal tracking errors, the observed pattern of abnormal returns associated with changes in the index may represent an equilibrium that is required to bring forth the required liquidity at the close on a change day. Whether indexers will continue to receive payments or will have to pay for liquidity depends upon the costs of providing the liquidity and the demand for that liquidity.

## APPENDIX

### Replication Model

To demonstrate the extreme sensitivity of tracking errors to slight deviations from index weights, consider a world in which there are 500 stocks with the same market values as those of the 500 stocks in the S&P 500 as of the end of 2000. In this world, assume that the function that generates the returns of each stock is statistically identical. By maintaining identical return-generating functions for each stock, we can concentrate on the effects of changing the weights of the component stocks.

The return-generating function of each stock is the one-factor model:

$$r = \pi + \varepsilon_i \quad (\text{A-1})$$

where  $\pi$  is a mean-zero factor common to all securities, and  $\varepsilon_i$  is a mean-zero independent disturbance for each security  $i$ . By selecting the variance of  $\pi$  to be 0.04 and the variance of  $\varepsilon$  to be 0.12, the common factor  $\pi$  will explain 25% of the variance of the total return.

In this model, the return for the S&P 500 is:

$$r_{500} = \pi + \overline{\varepsilon_{500}} \quad (\text{A-2})$$

where  $\overline{\varepsilon_{500}}$  is a weighted average of the  $\varepsilon$  for the 500 component stocks, and the weights are proportional to the market values of the actual stocks in the index as of the end of 2000. The



variance of  $r_{500}$  is the sum of the variance of the common factor and the variance of the weighted average of the unique disturbances. The resulting standard deviation of  $r_{500}$  is thus 20.28, 28 basis points higher than that of the common factor.

Similarly, the return of a portfolio invested in a subset of  $n$  of the 500 stocks in the index, where the weight given to each stock is proportional to its market value, is expressed in the same form as Equation (A-2). The difference is that  $\bar{\epsilon}_{500}$  is replaced with some other weighted average  $\bar{\epsilon}_n$ .

S&P has suggested that some indexers buy only the largest stocks. In this model, the standard deviation of a value-weighted portfolio of the largest 100 stocks in the index is 20.53, which is not much different from the 20.28 for the S&P 500 itself. The R-square is 0.9951, which is 1.00 to two decimal places. Yet the standard deviation of the tracking error is 145 basis points per year.

Dropping even a single stock can introduce meaningful tracking error. Consider dropping Nabor, whose market value as of the end of 2000 was the median of all 500 stocks. Specifically, Nabor is the largest stock of the smallest 250 stocks with a market value of \$8.7 billion, which represents just 0.07% of the market value of the 500 stocks in the index. If this one stock is dropped, and investments are maintained in the other 499 stocks in market proportions, the standard deviation of the portfolio return remains the same as that of the S&P 500 to four decimal places, but the standard deviation of the tracking error is 2.6 basis points.

## ENDNOTES

The authors thank for helpful comments and advice: Kenneth Kavajecz, Richard Kihlstrom, Craig MacKinlay, Ananth Madhavan, David Musto, Patrick O'Connor, Gus Sauter, Duane Whitney, Jeffrey Wurgler; and participants at the Wharton Finance Friday Workshop, the European Financial Association meetings, Virginia Polytechnic Institute and State University, Arizona State University, The Pennsylvania State University, and the University of North Carolina at Chapel Hill. Jianshan Xu provided outstanding research assistance. The authors gratefully acknowledge financial support from the Rodney L. White Center.

<sup>1</sup>U.S. publicly traded equity is the sum of corporate equities held by U.S. investors, equity issued in the U.S. by U.S. corporations and held by foreigners, and U.S. holdings of foreign equity, as estimated in "Flow of Funds Accounts of the United States" [2001, p. 82].

<sup>2</sup>The shares outstanding in Vestek are contemporaneous data. The shares outstanding in CRSP are sometimes adjusted to correspond to what CRSP considers the surviving company (with the benefit of hindsight). On occasion, the shares outstanding from the two sources differ substantially. Since the Vestek numbers are contemporaneous and those from CRSP are not, we use the Vestek numbers on shares outstanding to track more closely the shares S&P uses. To obtain shares outstanding within a month, we adjust the month-end shares out-

standing from Vestek using the share adjustment factors from CRSP. When there is an addition within a month, the month-end shares outstanding from Vestek for that month for that security are also adjusted, again using CRSP data.

<sup>3</sup>The opening price for Nasdaq stocks is the volume-weighted average price over the first 30 seconds after 9:30AM. In a limited number of cases, a stock begins trading only on or after the time at which it is added. In this case, we assume that it is purchased at the closing price on the change date, or, if this price is not available, at the opening price following the change date. The opening price for NYSE stocks is the price of the first reported trade on the NYSE.

<sup>4</sup>Let  $r_{sp}$  be the return on the index,  $r_{et}$  be the return on the early trading strategy, and  $r_{oa}$  be the return on any other asset. If  $b$  is the slope coefficient in a regression of  $r_{sp}$  on  $r_{et}$ , and  $\epsilon$  is the residual of that regression,  $\text{Cov}(r_{sp}, r_{oa}) = b \text{Cov}(r_{et}, r_{oa}) + \text{Cov}(\epsilon, r_{oa})$ . This last covariance is close to zero in the likely scenario that both  $\sigma(\epsilon)$  and  $\rho(\epsilon, r_{oa})$  are close to zero. If one assumes that this last covariance is 0, and since the estimate of  $b$  is 0.99956, the covariances using  $r_{et}$  are 0.04% higher than the covariances using  $r_{sp}$ .

<sup>5</sup>A possible criticism of our analysis is that the returns for the early trading strategy are based upon the shares outstanding from Vestek, which are not exactly the same as those that S&P uses. We have noted that S&P does not immediately adjust the number of shares outstanding for changes of less than 5%, but waits until the end of the quarter, while Vestek may adjust the number of shares it reports at other intervals. To determine the sensitivity of the incremental returns of the early trading strategy to this difference, we recalculate the S&P 500 index using the closing prices from CRSP and the month-end outstanding shares from Vestek adjusted forward or backward within a month using the share adjustments from CRSP. This recalculated series is the same as S&P would have calculated had it used these publicly available data. This recalculated series tracks the S&P 500 quite closely.

The monthly returns of the early trading strategy averaged 2.3 basis points more than the recalculated index, slightly higher than the 1.6 basis points reported in the text. The standard deviation of the monthly tracking error is 5.1 basis points, slightly lower than the 6.9 basis points reported in the text. Thus, the results stand, whether the returns on the early trading strategy are compared to the actual S&P 500 or to the recalculated index.

<sup>6</sup>Trading costs are capitalized into the value of a traded asset, so that the Barclays returns include trading costs.

<sup>7</sup>Many of these funds represent claims on the same portfolio of assets, and are really just different classes of stock that a mutual fund issues, with different fee structures.

<sup>8</sup>That most of these funds hold roughly 500 stocks does not in itself indicate that they follow an exact replication strategy, as that would require that the weights of the stocks in their portfolios equal the weights in the S&P 500 at all times. We could not match the stockholding information in Morningstar to the S&P holdings given in Vestek closely enough to allow the needed comparison.



<sup>9</sup>The close of trading on Nasdaq is 4:00PM Eastern time, with an allowance of 90 seconds for lags in reporting of trades, and the closing price is the price of the last reported trade. There are various time stamps associated with any trade. The closing prices reported on CRSP are not always the same as the price of the last trade reported on TAQ. Thus, we do not use the time stamp on TAQ to determine the closing price; we take the closing price for Nasdaq stocks from CRSP. NYSE closing prices come from TAQ.

<sup>10</sup>Analysis of volume at the close is based upon 161 additions and 46 deletions. Thus, of the previously reported 235 changes to the index, each involving an addition and a deletion, 74 additions and 189 deletions were dropped. The first pass eliminates any change involving: a bankruptcy, lack of data on the change day, an opening price on the day following the announcement of less than \$5.00, or more than 30 days between the announcement and the change. This pass eliminates 6 additions and 35 deletions.

Some changes to the index do not require any trading on the part of indexers, such as a merger of two S&P firms into a new S&P firm. For other changes, an indexer has a choice of what to trade, such as a merger of a non-S&P company and a S&P company into a new S&P company, as the indexer could trade in either to obtain or sell the shares on the new company. Eliminating these two types of changes reduces the number of additions analyzed by 49 and deletions by 154. This leaves 180 additions for analysis and 46 deletions.

Finally, in 2000, Nasdaq introduced after-hours trading, and because of the way the trades are reported, it is impossible to separate these trades from closing trades. This change affected 19 Nasdaq-listed stocks, and these were dropped.

The difference between the closing prices and reported better prices averaged 0.3% of the closing price, which is 15 cents on a 50-dollar stock. This average should be interpreted as an upper bound on the size of the compensation to indexers. It is an upper bound because of the possibility that the trade to the indexer involves two reported trades. A market maker may have agreed to buy the shares from a hedge fund at 0.45% less than the closing price and to sell these shares at 0.15% less than the closing price. The average discount is 0.3%, but the payment to the indexer is just 0.15%.

<sup>11</sup>The average daily percentage of the closing volume of NYSE stocks in the S&P 500 reported on non-NYSE markets is 7.6% from 1995 through 2000. This average is calculated, first, by averaging for each trading day the percentage of closing volume for NYSE stocks in the S&P 500 that is reported on non-NYSE markets and, second, by averaging these daily averages across days.

<sup>12</sup>Nasdaq-listed stocks account for much greater closing volumes than NYSE-listed stocks. These volumes should be interpreted with caution, as most trades for Nasdaq-listed stocks are principal trades and involve possible double-counting.

## REFERENCES

- Beneish, M., and R. Whaley. "An Anatomy of the S&P Game: The Effect of Changing the Rules." *Journal of Finance*, 51 (1996), pp. 1909-1930.
- . "S&P Index Replacements." *The Journal of Portfolio Management*, 29 (2002), pp. 51-60.
- Blouin, J., J. Raedy, and D. Shackelford. "The Impact of Capital Gains Taxes on Stock Price Reactions to S&P 500 Inclusion." Working paper, University of North Carolina, 2001.
- Bos, Roger. "Event Study: Quantifying the Effect of Being Added to an S&P Index." Standard & Poor's, 2000. Available at [spglobal.com/EventStudy.pdf](http://spglobal.com/EventStudy.pdf).
- Dhillon, U., and H. Johnson. "Changes in the Standard & Poor's 500 List." *Journal of Business*, 64 (1991), pp. 75-85.
- Edmister, R. O., A. S. Graham, and W. L. Pirie. "Excess Return of Index Replacement Stocks: Evidence of Liquidity and Substitutability." *Journal of Financial Research*, 17 (1994), pp. 333-346.
- "Flow of Funds Accounts of the United States: Flows and Outstandings." Board of Governors of the Federal Reserve System. Z.1 statistical release, June 8, 2001. Available at [federalreserve.gov/releases/Z1](http://federalreserve.gov/releases/Z1).
- Harris, L., and E. Gurel. "Price and Volume Effects Associated with Changes in the S&P 500." *Journal of Finance*, 41 (1986), pp. 815-829.
- Kaul, A., V. Mehrotra, and R. Morck. "Demand Curves for Stocks Do Slope Down." *Journal of Finance*, 55 (2000), pp. 893-912.
- Lynch, A. W., and R. R. Mendenhall. "New Evidence on Stock Price Effects Associated with Changes in the S&P 500 Index." *Journal of Business*, 70 (1997), pp. 351-383.
- S&P 500 2001 Directory*. Standard & Poor's, 2001.
- Shleifer, A. "Do Demand Curves for Stock Slope Down?" *Journal of Finance*, 41 (1986), pp. 579-590.
- Wurgler, J., and K. Zhuravskaya. "Does Arbitrage Flatten Demand Curves for Stocks?" *Journal of Business*, 75 (2002), pp. 583-608.

*To order reprints of this article, please contact Ajani Malik at [amalik@iijournals.com](mailto:amalik@iijournals.com) or 212-224-3205.*