

Investment Fundamentals: Benchmarking – Helpful & Painful

- Passive investment management and active management are philosophically different, in that the first is designed to closely replicate the performance of the broad market, as captured by an “index” like the S&P 500. Not a care is given to whether individual securities are overvalued or undervalued. Conversely, the goal of active management is to locate and purchase undervalued securities. Most US investors pursue active management in all or some portion of their portfolios, because nickels and dimes of better performance add up to significant sums over time.
- Interestingly, it has become the norm to “benchmark” actively managed portfolios against passive, or unmanaged, indexes. Choosing an appropriate index for benchmarking requires patience and attention to detail, because all indexes contain biases. Investors must be realistic about the limits and constraints in their own portfolios, and those in the chosen benchmarks.
- Additionally, active investors must decide how much deviation from a chosen index they will pursue and tolerate. To beat an index, the investors’ portfolios must deviate from the index.
- Indexes are a valuable tool for framing asset allocation decisions, measuring past performance, and defining future performance expectations. However, investing is not a science, so no tool should be used to the exclusion of all others, and no tool can take the place of thoughtful, well-informed critical thinking about the nature of investments and the goals of investing.

Claire B. Kendrick, Ph.D.
Amy Schondra Varvolis
Hari Pillai
March 2004

Hirtle, Callaghan & Co.

CHIEF INVESTMENT OFFICERS

Summary

According to *Webster's Collegiate Dictionary*, a “benchmark” is something that serves as a standard by which others may be measured. The dictionary goes on to describe an “index” as a number derived from a series of observations and used as an indicator or measure. In investment circles, a benchmark is an index that captures the behavior of certain financial instruments. The purpose of a benchmark is to give investors an indication of what they might expect from their investments in those financial instruments. For instance, the S&P 500, which is widely quoted, records the price changes for 500 of the largest capitalization US stocks. While not containing all the stocks in the US market, the S&P 500 is seen as a strong indicator of what occurs in the broad US stock market, and the many years of historical data for the index are seen as a plausible indicator of what investors might expect of current and future investments in that market. Thus, benchmarks are tools that can be used by investors for measurement and to manage expectations about the future.

In the 1960s, James Lorie and Lawrence Fischer of the University of Chicago compiled data on security returns for the Center for Research on Security Prices (CRSP). The CRSP index is generally recognized as one of the first well-studied and well-constructed indexes of the US stock market's performance¹. CRSP made available not only the aggregate index but also the information that went into constructing it, including returns on individual stocks, industrial sectors, and other groupings. Many other researchers have supplemented these findings and have worked

diligently to reconstruct histories of securities going back several hundred years. Beginning in the 1970s, a multitude of other indexes began to appear with the intent to capture the behavior of various asset classes and sub-classes across the globe.

The creation of the indexes led to the construction of portfolios that mimic the various indexes and can be bought by investors — index funds. The availability of index funds changed the investment landscape because investors could now “passively” invest or “actively” invest. The distinction between passive and active is significant. Passive investment means buying an index fund — a basket of securities that very closely tracks a specific index and is, therefore, focused on “relative” performance (relative to the index). Passive investment does not attempt to differentiate between securities that are overvalued or undervalued, because the goal is to track the index, not beat the index. On the other hand, active investment involves choosing individual stocks and bonds, or choosing managers that select individual stocks and bonds, based on some method of differentiating between securities that are overvalued and undervalued. Because each investor or manager has a method by which to select securities, there are inherent biases in the process and, so, biases in the portfolio. For this reason, active investment would have a more “absolute” performance orientation. However, over time, it has become the norm to compare actively managed portfolios with passive benchmarks as a tool to gauge performance. This comparison method, while valuable, is fraught with problems.

1 The Dow Jones Industrial Average, dating from 1896, was an early index, but its emphasis on tracking just a few stocks and its overweight to railroad issues prevented it from being widely used. The Cowles Index was one of the first high-quality, broad-based indexes of US stock returns. In the 1930s, Alfred Cowles gathered monthly capital appreciation and dividend yields on all NYSE-listed stocks from which a total returns index and a cumulative wealth index were constructed for the years 1872 through 1925. The S&P 500 Index began in 1928, but its prominence is more recent, because indexes and indexing were not a standard part of financial analysis until the 1970s.

For passive investment, benchmarking should not be an issue if a quality index fund has been purchased that accurately replicates the index. So, benchmarking (comparing actively managed portfolios to unmanaged indexes) is really an issue for those investors choosing active management. The disconnect between benchmarking (matching an index) and active management (beating an index) is rather obvious from the outset. In a nutshell, if the active management style too closely hugs the index, then why pay the higher management fee or take the time to pursue such a strategy when it is cheaper and quicker to engage a quality index fund? This is the dilemma: if investors want better returns, then they need to be willing to deviate from the benchmarks and recognize that there will be times when the active manager underperforms a chosen benchmark. Although there are financial practitioners that advocate both sides, Roger Ibbotson and Gary Brinson, in their book *Global Investing*, eloquently state

the situation by saying “it is rational and potentially very rewarding for active equity investors to continue the search for higher returns”.

In this paper, we begin with a discussion of passive and active management to clearly highlight the sharp philosophical differences. Next, we look at the explosion of indexes available to active investors seeking benchmarks, and the nuances or biases that investors can expect to find in all indexes. We also outline what features make a quality index. Finally, we note that indexes are a valuable tool for framing asset allocation decisions, measuring past performance, and defining future performance expectations. However, investing is not a science, so no tool should be used to the exclusion of all others, and no tool can take the place of thoughtful, well-informed critical thinking about the nature of investments and the goals of investing.

Passive or Active

How does one exploit the investment opportunities that exist across global asset classes? We begin with a very brief discussion of efficient and inefficient markets, because without a cursory knowledge of these theories, the difference between active investment management and passive investment management is blurred. We then examine passive management, because it forms a baseline against which active management is judged.

Market Efficiency or Not

Traditionally, investors have tried to achieve high returns by identifying and acquiring undervalued securities. Benjamin Graham and David Dodd are quite famous for providing rules for ascertaining value in their

books, *Security Analysis* written in 1934, and *The Intelligent Investor*, which was written in 1949 and has been published through several editions (latest 1986). Their methods for ascertaining value became the basis for traditional security analysis, which assumes that because of market inefficiencies, an asset’s market price fluctuates around its true value. Thus, the asset’s price is at times undervalued (a buying opportunity) and at times overvalued (a selling opportunity).

A second approach to investing rests upon the notion that markets are highly efficient and investors cannot ascertain which securities are overvalued and which are undervalued. In its most literal translation, the efficient market hypothesis says that the skills for identifying

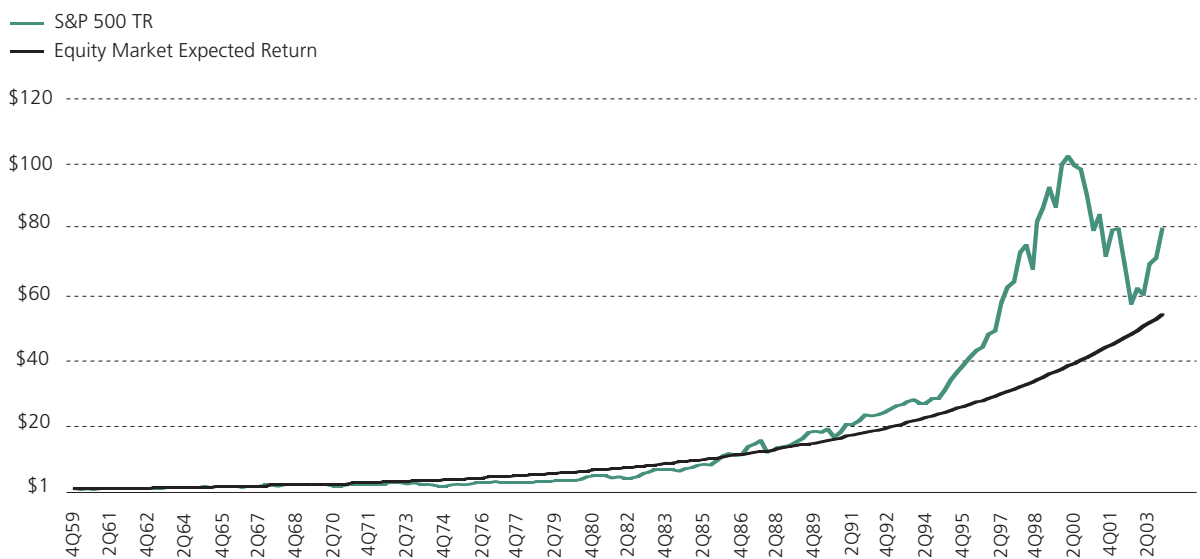
value cannot be acquired and consistently applied, so one investor cannot gain an advantage over another, except by pure luck. Instead of concern with overvalued and undervalued securities, investors should buy and hold a diversified set of securities.

Some of the fodder for efficient market theory was provided by examining the correlations of returns from one year to the next. The serial correlation of returns for US large capitalization stocks is practically zero, as calculated by Louis Bachelier in 1900 and reiterated by various statisticians since then. In other words, a strong year does not lead to another strong year or visa versa with any regularity. This led to a conclusion that annual stock returns are close to random (also known as a random walk). This may mean that no one can predict next year's returns by knowing those of this year. A random process, however, can produce results that visually resemble a trend. In the case of US stocks, this trend has been upward sloping (see **Diagram 1**) and has greatly affected investors' wealth. In **Diagram 1**, \$1 dollar invested in 1960 grew to over \$80 dollars by 2003, based on price appreciation and the re-investment

of dividends. The most assured method to capture the trend is to diversify one's holdings, which reduces the overall risk of investing. The risk declines as more and more securities are added to the portfolio, up to a point. Thus, investors should focus their effort on market analysis, which is the identifying of risks, expected returns, and correlations among the various investment opportunities or asset classes. Knowing what risks an investor can withstand and the investor's long-term investment needs, a specific mix of asset classes can be chosen to match those objectives — which is a major tenet of modern portfolio theory.

Many studies have shown modern portfolio theory to be consistent with the long histories of returns data. Few economists, financial managers, or investors question that there are gains from diversification or that risk taking can be well rewarded. However, practitioners do not agree that all securities are priced appropriately. Because capital markets are imperfect due to a host of factors, security prices may depart from their underlying values.

Diagram 1: Positive Slope in Stock Returns



Note: Equity Market Expected Return trend-line assumes a 9.5% expected return for US equity markets over the long run.
Sources: Ibbotson Associates and Hirtle, Callaghan & Co.

Numerous anomalies have been documented with respect to efficient market theory and modern portfolio theory, including the notion that investors are not rational in all their activities and make mistakes when it comes to valuing securities, and this irrationality itself causes securities to vary from their true value. Behavioral finance and behavioral economics, which look at the not-so-rational behavior of investors, are growing areas of research today². Other anomalies are outlined in **Diagram 2**.

Evidence of systematically mis-priced securities opens the door for a third approach to investing, which borrows the concepts of diversification and risk from modern portfolio theory, and pares them with an understanding of how security prices can vary from their true value, as embodied by traditional security analysis.

If indeed security markets are largely efficient, but not completely efficient, what is an investor to do? The answer is not simple and it is further frustrated by the fact that the various asset classes are not uniformly imperfect in their inefficiency. Studies have shown that the market for US large capitalization stocks is quite efficient, while the markets for smaller and non-US stocks are substantially less efficient. A gradation in inefficiency exists in the global bond markets also.

The existence of inefficiencies and the lack of uniform inefficiency can be handled through passive investment management or active investment management, but neither captures the totality of the securities markets. A mix of the two could be considered.

Passive Management

The goal of passive management is to track, not to beat, a specific index or performance benchmark. Passive managers do not try to identify bargain prices

Diagram 2: Pricing Anomalies

The principal anomalies where the observed stock price has periodically outperformed the market trend-line, include stocks that: (a) appear under-priced in terms of cash flows, earnings, or some other measure of corporate income, (b) offer dividends, (c) suffer neglect from security analysts or lack earning consensus by analysts, (d) issue earnings surprises, (e) have a small market capitalization (the “small cap effect” has shown that returns from the smallest quartile of publicly traded stocks have surpassed all other quartiles), (f) exhibit seasonality effects such as a jump in market prices in early January, near tax time, or coinciding with other major calendar events, or (g) are within the first few days following new issuance. The difficulty with these anomalies is that they appear perhaps more often than not, but cannot be predicted with the assurance to lead wise investors to place large wagers on them.

Source: *Global Investing*

or determine if a particular security is over-priced or under-priced. Instead, they employ indexing methods that will allow the portfolio (called an index fund) to replicate the benchmark returns as closely as possible.

Drawing on the philosophy behind the efficient market hypothesis, a few fund managers introduced index funds in the 1970s, but these investment vehicles did not gain wide popularity until the 1980s. Today, indexing is a popular choice in both stock and bond investing. Among US institutional investors (e.g., plan sponsors, pension funds, endowments, and mutual funds) with \$4.5 trillion in equity investments, the money placed in index funds accounted for \$1.2 trillion or 27% of the total, as shown **Diagram 3**. Data on how much non-institutional money is actively and passively

² In October 2002, Daniel Kahneman of Princeton University was awarded the Nobel Prize in Economics for having laid the foundation for a new field of research by discovering how human judgment may take shortcuts that systematically depart from basic principles of probability.

managed is hard to come by. However, for perspective, we note that US institutional investors with \$4.5 trillion represented just 44% of total US equity market capitalization, which stood at \$10.4 trillion in December 2002.

Popularity rests with the belief that index funds can provide returns equal to or superior to those of actively managed funds, and do so at a lower cost. Studies show that the broad indexes outperform the median active manager on average. Note that there are mathematical reasons for this. By construction, the aggregation of all active investors generates returns equal to the market. Thus, after transaction and management costs,

the average return of all active managers must be lower than that of the market. Even passively managed funds will have fees and costs associated with them, so that an index fund will generate actual returns below those of a hypothetical index. However, most actively managed funds have considerably higher fees and costs than those of passively managed funds. For a passive fund, fixed costs may be less, primarily due to lower personnel costs (little need for security analysts). The variable cost advantage of passive funds stems from lower commissions, because the turnover in the portfolio of index funds is generally much lower. Further, passive funds usually have significantly lower management fees.

Diagram 3: US Institutional Investment in US Equity Market

(\$ Billions)	Total Invested	Actively Managed	Indexed	% Indexed
Institutional Investors	\$2,145	\$1,250	\$895	42%
Mutual Funds	\$2,402	\$2,069	\$333	14%
Totals	\$4,547	\$3,319	\$1,228	27%

Note: Represents institutional ownership only (e.g., plan sponsors, pension funds, endowments, and mutual funds).

Does not include equity held by individuals, foreigners, governments, or insiders.

Source: S&P; data as of December 2002

Building an Index

The first step when building an index fund is to decide what index to follow. Second, a method must be selected for achieving performance that tracks the index as efficiently as possible. This sounds simple, but it is not. A method of “full replication” is expensive because an index itself can have substantial turnover and many securities in the index may be highly illiquid or small, so that they cannot be bought in the time frame or quantity as needed. This problem is particularly severe for the broadest indexes (such as the Wilshire 5000 and Russell 3000, which we discuss later) and the more inefficient markets (such as emerging market equities and high yield bonds), where full replication is not feasible. In order to reduce

the costs and offset the inability to buy all the requisite holdings, techniques such as sampling and optimization are employed.

Sampling involves selecting a subset of the index to hold. The goal is to choose a subset that will mimic the entire index with minimal tracking error. For equities, two popular methods of sampling are size-stratified sampling and industry sampling. In size-stratified sampling, the companies in the index are divided into size categories and stocks are selected to represent a proportional capitalization weight. Industry sampling involves dividing the index into industry groups and replicating the same industry distribution in the portfolio. For bonds, sampling is based on duration and credit quality. A more complex technique for tracking

an index is optimization. For a given number of securities (selected from the universe of securities, not just those in the index), the optimizer selects those that will create minimum tracking error. As the number of securities selected is allowed to increase, the tracking error continues to drop.

Third, the weighting method of the index must be decided. The most frequently used portfolio weighting methods are (a) by market capitalization or (b) market float. For indexes of foreign stocks or bonds, there is also the issue of securities of one country versus another, which requires an overlay, usually by GDP or GNP, on the market cap or market float.

Market capitalization is the most popular method because it requires the least rebalancing, since a security's value and its weight in a cap weighted index move in lockstep. (Remember that capitalization weighting is simply price times the number of shares outstanding, so that as price moves up or down, the cap weight moves up or down.) Rebalancing is only required for composition changes (companies entering and exiting the index) and dividend re-investment. Determining float-adjusted weightings is more labor intensive. Float refers to the number of shares that an investor can conceivably purchase, so that family holdings, corporate cross-holdings (one company's significant ownership in another), government holdings, ESOP shares, and other closely held shares are subtracted from the number of shares outstanding. For instance, UPS, the package delivery company, is largely owned by its employees, so that the number of shares available for purchase by non-employees is practically non-existent. As such, UPS may not show up or may show up as a small cap company in a float-adjusted index, but it would be a large cap stock in a capitalization-weighted index. Similarly, Avis, United Airlines, and countless

other large well-known companies with significant employee ownership would appear differently among float-adjusted indexes compared with market capital weighted indexes.

Active Management

Actively managed funds attempt to provide the investor with added value through the selection of stocks that appear to be undervalued, so the resulting portfolio may look very different from that of an index fund. The magnitude of the deviation from the index will reflect the particular style of the investment manager. Active management is prefaced on the conviction that market conditions periodically offer the opportunity to enhance the performance of the portfolio. Thus, the choice between passive and active management of a portfolio involves two basic issues:

- the investor's opinion on the efficiency of markets, e.g., whether opportunities to earn excess returns from active management exist, and
- the perceived ability of a given investor to recognize and exploit such opportunities, if they exist, e.g., locate, select and monitor active managers.

While the efficient market hypothesis has led to many useful analytical tools, the underlying assumption about the inability to recognize mis-priced assets does not always hold true in the global environment. In fact, it is the complexities of the global market that often produce price anomalies and provide a strong basis or opportunity for active management (see **Diagram 2**).

However, it is not enough that a security is found to be under-priced. Knowing or estimating the magnitude of the mis-pricing is also critical for profitable active management. Since assets and asset classes represent

claims on future income, their values can be determined by summing the present value of future cash flows. In an uncertain world, a wise investment choice will be based upon a projection of future cash flows and the projection will need to exceed the equilibrium discount rate, which includes three components:

- the real riskless rate (uniform for every asset class, usually US Treasury bills),
- an expected inflation premium (uniform for every asset class), and
- a risk premium (specific to each asset class).

Investment would not be warranted if investors are not rewarded for taking investment risks (i.e., the returns or expected future cash flows do not meet or exceed the three components of the equilibrium discount rate). While this sounds straightforward, it is not. The task of determining the asset class-specific risk premium requires the analysis of potential cash flows, which is complicated by changes and innovations in the capital markets, uncertainty about national and global economic strength, currency fluctuations, technical substitution and obsolescence, management risk, policy and legislative changes, wars, etc. Moreover, inflation can be very volatile, so the inflation premium can be difficult to pin down. While obviously troublesome for the analyst and active fund manager, these events also provide exploitable opportunities for those willing to incur the costs and time to obtain the pertinent information.

The Active Investor

For active investors, there are two critical considerations:

- (1) choosing the right stocks and bonds individually, or choosing the right money manager who will select the underlying stocks and bonds, and

- (2) choosing a benchmark against which to compare the active holdings.

Benchmarking is more art than science, because there is a significant philosophical dilemma right from the start. The philosophical conundrum is that benchmarking is designed to measure how closely a portfolio mimics a chosen index, while active management is the desire to beat or surpass the returns of a chosen index. How can benchmarking be done satisfactorily? It cannot be. Active investors must be willing to deviate from the benchmark in order to beat the benchmark. After all, there is no point in paying for active management (higher fees and trading costs), if the portfolio is going to closely match that of an index and its related index funds.

Why should active investors use benchmarks? The answer is that benchmarking is a valuable, although less-than-perfect, tool in terms of measurement and shaping future return expectations. It is somewhat like a discussion of US-style democracy, which asks why any country would want the messy process of electing officials at the local, state, and national levels along with referendums, filibusters, and the other antics that ensue. The answer is that despite its messiness and less-than-ideal characteristics, it is still the best the world has when it comes to representation for the people and by the people.

Benchmarking is messy because it requires subjective judgment as to (a) how much variation from a benchmark investors want or can withstand in each asset class, and (b) by how much they want their overall portfolios to deviate from the mix of asset class benchmarks (called a blended benchmark). Active management is not for the faint-hearted because it requires steps (1) and (2) (see paragraphs above), as well as steps (a) and (b) listed in the previous sentence. Who should and who should not pursue active management is a valuable discussion, but is beyond the scope of this

paper. Warren Buffet is one of the more famous active investors and he is not alone, since at least some active management is undertaken by the majority of investors in the US.

In the next section, we aim to highlight some additional messiness, which is that all benchmarks

have quirks. All active investors have limits and constraints and, therefore, biases embedded in their investment choices; also, every benchmark has its own biases. Thus, no benchmark is a good fit for all active investors and, in fact, no benchmark is a perfect fit for any investor.

Benchmark Nuances

An abundance of market indexes are available today. There are indexes for the primary asset classes (large capitalization stocks, small cap stocks, foreign stocks, US high grade bonds, US high yield bonds, foreign bonds, etc.), as shown in **Diagram 4**. Indexes also abound for sub-classes such as value stocks versus growth stocks. Further slicing and dicing can be done down to painstaking detail, and if no publicly available index can be found to match the investor's undertakings, then an index manufacturer can be engaged to develop and maintain a custom-made benchmark.

For the sake of simplicity, we discuss four of the most well-known and well-publicized benchmarks, each that has the purported goal of measuring the broad US stock market and shedding light on the expected current and future performance of investments in the broad US stock market. The four indexes chosen for our review are the S&P 500, the Dow Jones Industrial Average (DJIA), the Russell 3000, and the Wilshire 5000.

This comparison illustrates the nuances and variations among indexes in general, particularly the differences in the construction methodologies. It is not meant to expose every quirk. In essence, each index has biases and shortcomings. Interestingly, the broad US stock market, which is heavily dominated (around 75%) by

large capitalization stocks, is the easiest market in the world for which to build indexes and index funds. When classes, such as high yield bonds or emerging market stocks are considered, the disparities among the competing indexes are found to be at least five times greater than those found among the broad US market indexes.

S&P 500

The S&P 500 is the most widely used index. Despite an emphasis on large cap stocks, it includes a significant portion of the total value of the market (nearly 80%), and is thought to well represent the broad market. The S&P 500 is a market capitalization weighted index, meaning that each company's stock price is multiplied by its shares outstanding. The largest stocks in the index, therefore, will account for a large portion of the index's return using this method. Conversely, the smallest stocks in the index will have very little effect on performance.

Instead of a highly systematic selection method, a committee at S&P chooses stocks for addition or deletion based on highly subjective criteria. Companies elected are usually US-based, but the US exchange-traded ADRs (American Depositary Receipts) of several non-US companies are also included. Candidates are expected to have reasonable liquidity

Diagram 4: Commonly Used Indexes

Equity Indexes	
US	Non-US
<p>Large Cap & Broad Market</p> <p>AMEX Composite DJIA NASDAQ Composite NYSE Composite Russell 1000 Russell 3000 S&P 500 Wilshire 5000</p>	<p>Developed</p> <p>MSCI EAFE FTSE All-World Developed Nikkei</p>
<p>Medium & Small Cap</p> <p>Russell 2000 Russell 2500 S&P Midcap 400 Wilshire 4500</p>	<p>Emerging</p> <p>MSCI EMF IFC Investable</p>
	<p>Global</p> <p>FTSE All-World MSCI World MSCIAC World DJ Global SSB Global Equity</p>
Fixed Income Indexes	
<p>Investment Grade</p> <p>Lehman Aggregate Lehman Credit (Corporate) Lehman Government Lehman Municipal Bond</p>	<p>High Yield</p> <p>CSFB Global HY Lehman Brothers HY Merrill Lynch HY Master Salomon HY Market</p>
Alternative Investment Indexes	
<p>Private Equity</p> <p>Cambridge Associates NACREF NAREIT Venture Economics</p>	<p>Hedge Funds</p> <p>Altvest CSFB Tremont EACM Hedge Fund .Net Hennessee Group HFR LJH Global Investments MAR Van Hedge</p>

in their shares (e.g., the ratio of the annual dollar value traded to the market capitalization should be 0.3 or greater), and each should have a market capitalization of \$4 billion or more. The market cap criterion has shifted over time. Companies must show financial

viability, meaning four consecutive quarters of positive reported earnings. There must be public float of at least 50% of the stock (at least 50% of the outstanding shares of the company must be available for the public to trade); however, the index weights of each company are

based on the total number of shares outstanding, not the float. Finally, each company must be an operating company, not a closed-end fund, holding company, investment vehicle, royalty trust, or partnership (exception: REITs are eligible). A company is deleted from the S&P 500 if it is significantly restructured and no longer meets the inclusion criteria. Some of the salient features of the S&P 500 are shown in **Diagram 5**. The table specifies just some of the complexity behind the construction of the S&P 500 and the other three indexes, and is not meant to be exhaustive.

S&P believes that unnecessary and excessive turnover in index membership should be avoided when possible; therefore, the index committee will allow a company

to remain in the index if it believes that the company has only temporarily violated one or more of the inclusion criteria. Further, although many companies may meet the selection criteria at any one time, they are not admitted until an existing member of the index has violated criteria and is removed. In other words, reconstruction or change to the holdings in the index is executed as needed, not on a scheduled basis, which is a philosophical difference compared with the Wilshire 5000 and the Russell 3000. The attraction of the S&P 500 is its broad capture of the market (nearly 80%) with relatively fewer companies, and its low turnover. These features make it relatively easy for index funds to mimic, hence the S&P 500 is the most often replicated index in the world.

Diagram 5: Index Characteristics

Index	S&P 500	DJIA	Russell 3000	Wilshire 5000
Year Established	1928	1896	1984	1970
Measurement Objective	Broad Market	Broad Market	Broad Market	Broad Market
% of Market Represented	78%	24%	98%	100%
Number of Companies	500	30	3,000	5,200+
Distribution by Market Cap				
NYSE	83.5%	95.8%	79.6%	78.0%
AMEX	0.2%	0.0%	0.4%	0.6%
NASDAQ	16.3%	4.2%	20.0%	21.4%
Criteria for Inclusion	Subjective — S&P selection committee; stocks chosen for market size, liquidity, and industry group representation	Subjective— WSJ editor selection committee; representative of major US industry sectors	Objective — Top 3,000 NASDAQ, NYSE, AMEX stocks; ranked by market cap but weighted by float	Objective — All US equities headquartered in the US, for which prices are available on May 31 each year
Weighting	Market Capitalization	Price	Market Capitalization; adjusted for large private holdings, and corporate cross ownership	Market Capitalization
Float Adjusted	No	No	Yes	No
Reconstitution	As Needed	As Needed	Annually	Monthly

Source: Frank Russell Company, SSGA, Wilshire Associates

DJIA

The Dow Jones Industrial Average (DJIA), or the Dow, is another widely quoted index, and is one of the oldest broadly published market indicators. The origins of the index date back to the late 1800s when even stockbrokers on Wall Street found it difficult to discern from the daily jumble, of up-a-quarter and down-an-eighth, whether stocks generally were rising, falling, or treading water. Charles Dow devised his stock average (comprised of 11 stocks, mostly railroads) in 1884 to attempt to make sense out of the confusion.

Today, the DJIA is comprised of 30 large cap stocks from various industries determined by its creators to be “bellwethers of the US stock market”. Unlike the other indices reviewed here, the DJIA is a price weighted index, not a capitalization weighted index. This means that the weightings of the companies within the index are affected only by changes in the stocks’ prices. The number of shares outstanding and those that are available to the public have no impact on the composition or performance of this index. When the DJIA was initially created, its value was calculated by simply adding up the stock prices of the constituent companies, and dividing by the number of companies. However, this proved somewhat problematic as stock splits and other corporate actions began to have a noticeable effect on the index. Later, the practice of adjusting the divisor was introduced to smooth the effects of these events. Because the changes in the denominator are complex, there are few price weighted indexes.

Inclusion rules for the DJIA are the least “scientific” or systematic of the indexes presented here. Members of this index are selected at the discretion of a group of editors of *The Wall Street Journal*. There are no pre-determined criteria except that candidates are

established US companies that are leaders in their industries. Generally, a stock is added if it is deemed to have an excellent reputation, demonstrates sustained growth, is of interest to a large number of investors, and accurately represents the sector(s) covered by the average. Composition changes are rare, and generally occur only after corporate acquisitions or some dramatic shift in a company’s core business. When this occurs, the entire index is reviewed, and there may or may not be changes to the overall index. While the DJIA only explicitly represents 24% of the broad US stock market, the liquidity or high tradability of the shares is perhaps the greatest of the four indexes — providing the clearest indication of daily market activity. Because the price movements of the stocks cause the individual weightings to change constantly, the burden to index funds trying to replicate the index is high.

Russell 3000

The Russell 3000 Index was created in 1984 and is actually a combination of two sub-indexes, the small-cap Russell 2000 and the large cap Russell 1000. The Russell 3000 holds 3,000 companies, representing 98% of the investable US equity market. The smallest 2% of companies are purposefully excluded from the Russell 3000 Index, as there is so little liquidity available in such small companies. The Russell Indexes are designed to facilitate the building of index funds. As such, there are Russell Indexes that emulate large pieces of the market, ones to capture different styles (value or growth), and others embody a myriad of other characteristics.

Unlike the S&P 500 and the DJIA, the Russell 3000 has more structured rules of inclusion. For annual membership, all US stocks are ranked from largest to smallest by market capitalization as of May 31. Determining a company’s total market cap requires combining all outstanding shares. However, the weight

assigned to each stock in the index is based upon the float, not the total market cap. As such, companies are weighted by available market capitalization, after adjusting for cross-ownership and large private holdings. Large private holdings are defined as individual, corporate, or trust holdings greater than or equal to 10% of shares outstanding. Simply, float-adjustment reduces the weightings to reflect what the average investor could conceivably buy. For companies with more than one series of stock available, the weighting will reflect the price and the number of float-adjusted shares of the largest or most liquid series (computations that can be difficult). The top 3,000 stocks incorporated in the US and its territories become the Russell 3000 Index effective June 30.

All stocks listed on the NYSE, AMSE, or NASDAQ are considered for inclusion. Securities not eligible include stocks trading below \$1.00, closed-end mutual funds, limited partnerships, royalty trusts, non-US corporations, foreign stocks, and ADRs. A stock is removed from the index prior to the next reconstitution if the company is de-listed. Russell does not replace stocks that have been deleted from the index between reconstitution dates. Additions to the index, except during the annual reconstitution, only occur because of spin-offs. To be an eligible spin-off, the company's total market capitalization must be greater than the float-adjusted total market capitalization of the smallest security in the Russell 3000 Index at its last reconstitution.

The annual reconstitution is a distinguishing feature. Russell believes that complete reconstitution is important since market conditions and characteristics change over time. While passive taxable investors may be concerned with the consequences of turnover in their index funds due to reconstitution, Russell

is of the opinion that annual reconstitution results in the most effective balance between turnover and market representation.

The scheduled reconstitution and more objective inclusion rules are a positive for investors interested in clarity about market activity over the medium to longer-term. However, while the scheduled reconstitution allows index funds to prepare for the changes, the change can affect 10% or more of the shares. This huge and costly annual dislocation may discourage index manufacturers from choosing the Russell 3000 over the S&P 500. Also, the illiquidity of the smallest capitalization stocks and the large number of companies may make it difficult for all but the largest index funds to closely mimic the Russell 3000.

Wilshire 5000

The Wilshire 5000 is one of the broadest measures of the US equity markets. It reflects the performance of all US headquartered equity securities with readily available price data. When the index was created in 1974, it included approximately 5,000 stocks, hence its name. However, the index held more than 5,200 companies on December 2003, and comprised more than 7,000 companies prior to the bursting of the tech bubble in 2000.

For inclusion in the Wilshire 5000, a company must be headquartered in the US. Issues may include common stocks, REITs, and limited partnerships. Companies excluded from the index include "bulletin board" issues (e.g., penny stocks that do not trade very often, also called pink sheet stocks), or any company that is deemed not to have readily available pricing. Additions to the index are made once a month after the month-end close. IPOs are added at the end of each month, and spin-offs are added as soon

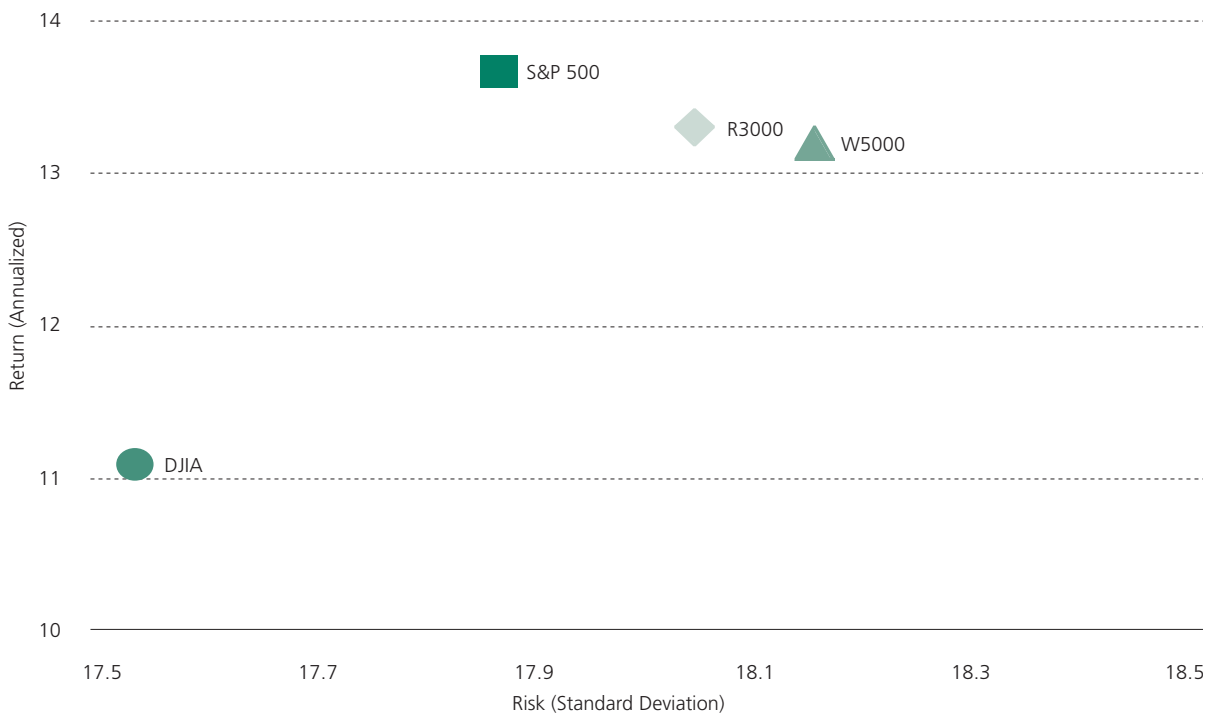
as is prudently possible. A stock which stops trading will remain in the index for two months, either at its last exchange traded price or, if obtainable, at its first available pink sheet price. At the end of second month, if the stock has not yet resumed exchange trading, it will be dropped from the index.

While the index provides a good approximation of dollar changes in the US equity market, it may not be the best representation of the investable market, as the smallest companies in the index (the bottom 2%, or approximately 2,400 companies) have little to no liquidity. The timeliness with which market changes are incorporated is appealing to investors wanting near-term market clarity. However, the combination of many small illiquid issues and the rapid reconstitution make it difficult for index funds to replicate the Wilshire 5000.

Comments on Comparability

In sum, the S&P 500 will appeal to those seeking a popular, low-cost, large capitalization orientation. Investors that emphasize daily liquidity in their investment thesis will like the DJIA. The Russell 3000 is best suited to investors who desire a broad equity market exposure and highly systematic construction methodology. The Wilshire 5000 is designed for those who consider breadth a critical feature of a benchmark. Each index has unique strengths in terms of its ability to relay to investors the changing scope of the market. Each index also has weaknesses because no stock selection method is without some bias. In **Diagram 6**, the historical risks and returns for each of the four indexes are displayed.

Diagram 6: Risk and Return Differences (1980-2003)



Source: Ibbotson Associates

It is worth noting that much of the variation among the four indexes stems from philosophical differences, not logistical issues. Perhaps the most influential difference is the portion of holdings found in large cap versus mid cap versus small cap stocks. When small cap stocks outperform, it is reasonable to expect the Wilshire 5000 and the Russell 3000 to produce returns that exceed those of the S&P 500, and certainly that of the DJIA. Conversely, strong performance by large cap stocks will advantage the DJIA and the S&P 500.

Which index is the better indicator of the broad US market? If investors cannot actually buy the broad US market, is it meaningful for investors to use such an index as a benchmark for their more limited and constrained portfolios? Does an investor benefit more from a benchmark tailored specifically to match the limitations and constraints in the portfolio, or is such behavior self-serving and not instructive? These questions form the basis for the next section.

Characteristics of a Quality Index

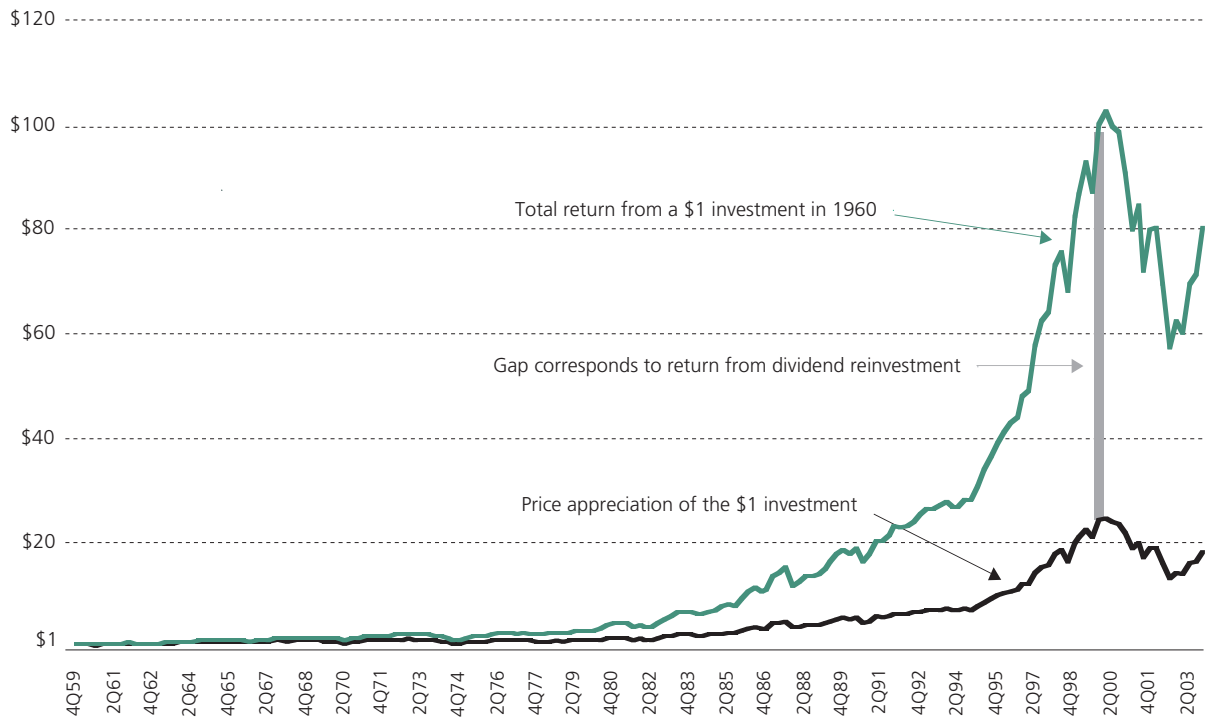
There is no shortage of indexes. In this section, we discuss the general characteristics of a quality index. Like all measurement tools, the quality of an index depends in large part on how the index is intended to be used. As we showed previously, the S&P 500 is calculated in a very different fashion from the DJIA, the Russell 3000, or the Wilshire 5000. The purpose of these indexes is to measure broad US stock market performance. Since each of the four has been built with different criteria, each may be useful in its own way. According to Roger Ibbotson and Gary Brinson, two highly skilled practitioners, four criteria distinguish a quality index.

- (1) It should reflect what an investor facing no transaction costs could achieve in the chosen underlying market.
- (2) It should be reflective of a broad class of securities, either by containing all of the securities in the class or by careful section of a meaningful subset.
- (3) It should have a long-dated history.
- (4) It should be updated in a timely fashion in order to closely track the underlying market that it measures.

If an index is to be a benchmark for portfolio performance, an index needs to be tradable, as specified in criteria (1). In essence, an index for which an index fund cannot be constructed is generally a poor index. Further, Ibbotson and Brinson recommend that the index be computed on a “total returns” basis, which means that not only is capital appreciation measured but also dividend income. Dividend income that is reinvested can add materially to an investor’s wealth over the long run, as shown in **Diagram 7**.

If an index is not meant to track a broad market, then there is really no need for it (criteria (2)). Generally, a highly idiosyncratic investment style is not looking for validation by comparison with a structured index. Indexes may be broad or narrow in what they contain, but the purpose should be to emulate a broad market. For instance, although the DJIA is a narrow index by construction (just 30 stocks), it is designed to capture the broad US stock market, or at least the US large cap market. The DJIA is recalculated every 30 seconds. This is meaningful because in any given 30-second period most or all of the 30 DJIA stocks will have traded.

Diagram 7: Growth of \$1 Investment



Source: Ibbotson Associates and Hirtle, Callaghan & Co.

A 30-second recalculation of the Wilshire 5000 would be pointless, since the smallest of the 5,000+ stocks would not likely have traded. However, the Wilshire 5000 may provide a stronger indication of long-term changes in investors' wealth than does the narrowly defined DJIA. Both indexes have their place.

A long history is a useful property, as stated in criteria (3). While most commonly scrutinized for their return expectation, indexes should also accurately supply information to investors about volatility and correlations with other asset classes or events. Without a history that spans diverse economic scenarios, this comparability function is compromised. For exactly that reason, when launching an index, modern index manufacturers often reconstruct an historical record.

Finally, with regard to criteria (4), since investors are primarily concerned with history's ability to illuminate current and future outcomes, the index should be updated in a timely fashion so that the insights that the index provide can be put to use at all times. This timeliness should include dividend income recorded in the period received.

While these four criteria define a quality index, what are the considerations most often cited by the index manufacturers themselves? As expected, index manufacturers are not without bias in favor of their own indexes and their views about the best traits in an index have influenced the expectations of investors. Investors are encouraged to think critically about the attributes of an index in light of their own investment goals and distinguish between quality and marketing hype.

A Manufacturing View

Those firms that track public securities and create benchmarks for these securities, whom we refer to as benchmark manufacturers, cite four considerations that investors should embrace:

- (1) Turnover
- (2) Trading costs
- (3) Tracking objective
- (4) Popularity

First and foremost, we need to establish that most manufacturers want to sell investors a passively managed portfolio, not be the measuring rod for their actively managed portfolios. There is precious little money in selling the returns data, particularly in the age of the internet where information, both public and proprietary, is easily tapped by millions.

Looking at the list, consideration (1) is turnover or how quickly company positions are bought and sold. Remember that there are deep philosophical differences among the manufacturers. The S&P 500 opts to reconstitute itself very infrequently, because it believes that is more in tune with how investors would handle their own portfolios. Perhaps more to the point, limited turnover makes it easier to construct and maintain an index. On the other side of the philosophical divide is the Wilshire 5000, which reconstitutes each month. Wilshire believes that investors should have clarity on the full dynamics of the marketplace, even though it makes an index more difficult to construct. This leads easily into consideration (2), which is trading costs.

Because indexes designed to match the Wilshire 5000 and the Russell 3000 have many more company holdings, are more complicated to construct, and are more difficult to reconstitute, they will generally have higher trading costs. A study by State Street Global

Advisors (SSgA) in recent years examined trading costs and found stunning differences, with small cap costs per share averaging three to four times higher than those for large cap shares. If the higher trading costs are netted from the returns of the index fund, then indexes with less frequent reconstitution and fewer company positions may appear to be better performers. Better performers in which market environments, over what time frame, and at what risk level? Refer back to **Diagram 6**.

Consideration (3) recognizes that index funds are designed to track the index with as few company holdings as possible. The fewer the holdings, the lower the trading costs. When the number of holdings are significantly less than those in the chosen index, there will be market environments when the index fund notably deviates from the index. Likewise, actively managed portfolios will deviate from their benchmarks significantly under certain market scenarios. We discussed that active investors must expect these deviations and decide which deviations they will tolerate over the short, medium, and long run. Index sensitivity among active investors has grown to the point where many highly regarded “active” managers (true stock pickers) have a 95% correlation to the chosen indexes. Is this desirable? No, it means that (a) investors are paying active management fees for what are really “closet” index funds and (b) investors have little chance to significantly beat the index over a market cycle, because without variation there exists no opportunity for better performance. To beat an index, one must vary from it.

Popularity, consideration (4), should not be a concern for active investors seeking a benchmark. It may have some bearing for passive investors searching for an index fund, because popularity may lead to larger index

funds and greater liquidity, but even that may be moot beyond a certain minimum size. In the US index market, the S&P 500 is the undisputed winner, as shown in **Diagram 8**. Nonetheless, that does not make the S&P 500 the undisputed winner with regard to benchmarking — its limited exposure to small cap stocks means that investors tracking the S&P 500 are encouraged to limit their holdings in the highest return generating asset class. (Remember the “small

cap effect”, in which studies have shown that returns from the smallest quartile of publicly traded stocks have surpassed those of all other quartiles.)

The net result of the marketing and relentless growth in index products is that many investors have lost sight of the original intent of benchmarking — it is one method for measuring performance. Matching the benchmark is not an end in itself, for a whole host of reasons, many of which we have discussed at length in this paper.

Diagram 8: Popularity of Major US Indexes

\$ Billions	Total Market Capitalization	Passively Managed Assets	Passively Managed Assets %
S&P 500	\$8,105	\$626	7.72%
Russell 3000	\$9,976	\$69	0.69%
Wilshire 5000	\$10,260	\$73	0.71%

Source: SSgA, Goldman Sachs, Fact Set; data as of June 2003

Conclusion

Investors choosing passive investment management need to concern themselves with selecting the right index fund or funds to achieve their investment goals. However, once this is completed, much of their work is over. On the other hand, for investors electing active management there are multiple activities to undertake.

- (1) Choose the right money manager who will select the underlying stocks and bonds.
- (2) Choose a benchmark against which to compare the active holdings.
- (3) Decide how much variation from a benchmark is acceptable to them in each asset class.
- (4) Decide how much to allow the overall portfolio to deviate from a combined mix of various asset class benchmarks (called a blended benchmark).

We discussed at length some of the problems an active investor can expect when trying to select an index, given the quirks and nuances in their own portfolios and those in the indexes themselves. Investors need to acknowledge the strengths and shortcomings in their own portfolio construction and those in the benchmark indexes. With an open-minded approach, investors should be able to choose a benchmark and press to overcome the inherent philosophical mismatch between tracking an index and beating an index — admitting that the benchmark is less-than-ideal but still valuable when it comes to portfolio analysis and managing expectations about potential future returns. Investment success is not matching the index or indexes chosen, because if that is the goal, then the investor should just switch to passive management.

Investment success is also not confined to just beating the index or indexes. The question logically arises as to how much risk was taken to beat the index. If the investor is not comfortable with that level of risk day-in and day-out, then the solution is to dial down the risk/return expectation, not dial up the risk quotient just to beat some less-than-perfect measurement system. Also, investors need to ask themselves, over what time horizon do they expect to beat the benchmark. The historical average return of a benchmark is just that — an average. The true volatility of returns within that index should be explored. An active portfolio will likewise show volatility, often times more volatility than the chosen index. Thus, it is unrealistic to believe that it is possible to consistently beat the benchmark. A more realistic expectation is to beat the benchmark over longer-term horizons, such as five or ten year periods. (Please refer to our previous paper of June 2003 entitled “*Investment Fundamentals: Investment Horizons vs. Planning Horizons*”.)

If investment success is not matching the index and investment success is a broader concept than just beating an index, what else should an investor consider? Although the list of other considerations will vary from investor to investor, a simple list might include four things.

One, all investors must clearly delineate their own investment goals. For instance, the investment goals of a family striving to send children to college and retire at age 55 will be very different from those of GM's pension fund, which has to meet the health and income obligations of retired employees, when the number of retired individuals exceeds the number currently working.

Two, investors should require that their chosen investment manager or managers have clearly articulated their strategies. From this, investors need to make a determination that these strategies continue to dovetail with those of their own over time. Investors should seek to have confidence in their managers' ability to navigate through the market pitfalls and stay in sync with the investment mandate.

Three, investors need to familiarize themselves with the methodologies and construction practices of any benchmarks they choose for measurement purposes. A lack of understanding of the nuances can lead to unrealistic and unrealized expectations for returns.

Four, investors should have established a schedule for reviewing their long-term investment returns with respect to their long-term needs and expectations (e.g., not get overly excited about the short-term volatility in their portfolios or their chosen benchmarks). Realistic expectations are important, because otherwise the investment experience will be frustrating and demoralizing even if the managers have clearly met their mandates and the benchmarks chosen are the most appropriate ones.

In sum, active management is pursued by most investors in at least some portion of their portfolios. Benchmarking, which is a less-than-perfect tool for measuring investment performance, is nonetheless valuable because no perfect tool exists. However, no tool can take the place of thoughtful, well-informed critical thinking about the nature of investments and the goals of investing.

Investment Fundamentals: Benchmarking – Helpful & Painful is one issue in a series of articles relating to general financial matters prepared by Hirtle, Callaghan & Co. from time to time and for the exclusive use of its clients. While Hirtle Callaghan believes this information may be of interest to you, it should not be construed as legal or tax advice, nor is it a substitute for careful tax planning. Also, keep in mind that future legislative or administrative changes or court decisions may materially affect the tax consequences of your investments.

Information herein is not intended to be a complete analysis of every material fact respecting any company, industry, or security in our portfolios. The opinions expressed herein reflect the judgement of Hirtle, Callaghan & Co. at this date and are subject to change without notice. The securities identified may or may not be held in client accounts, and this presentation is not intended to be, and should not be construed as a recommendation, offer, or solicitation with respect to the purchase or sale of any security.

© 2004 Hirtle, Callaghan & Co.

Hirtle, Callaghan & Co. — Vision. Integrity. Diligence.

In 1988, Hirtle Callaghan introduced a new standard of service to the investment marketplace. That service is based on the complete investment objectivity and active advocacy of a professional Chief Investment Officer, supported by a fully staffed Investment Department. This time-tested approach leads to fundamentally superior investment programs and represents the only proven way to ensure better net performance and greater fiduciary peace of mind in perpetuity. Today, Hirtle Callaghan is America's leading Chief Investment Officer organization.

If you would like to discuss how Hirtle Callaghan can develop an ideal investment program to capture better net performance and greater peace of mind for your family or organization, please contact us at (610) 828-7200 or www.hirtlecallaghan.com .

Investment Fundamentals is a quarterly publication of Hirtle, Callaghan & Co.

Hirtle, Callaghan & Co.

CHIEF INVESTMENT OFFICERS