The Efficient Market Hypothesis and Its Critics

Burton G. Malkiel

A generation ago, the efficient market hypothesis was widely accepted by academic financial economists; for example, see Eugene Fama’s (1970) influential survey article, “Efficient Capital Markets.” It was generally believed that securities markets were extremely efficient in reflecting information about individual stocks and about the stock market as a whole. The accepted view was that when information arises, the news spreads very quickly and is incorporated into the prices of securities without delay. Thus, neither technical analysis, which is the study of past stock prices in an attempt to predict future prices, nor even fundamental analysis, which is the analysis of financial information such as company earnings and asset values to help investors select “undervalued” stocks, would enable an investor to achieve returns greater than those that could be obtained by holding a randomly selected portfolio of individual stocks, at least not with comparable risk.

The efficient market hypothesis is associated with the idea of a “random walk,” which is a term loosely used in the finance literature to characterize a price series where all subsequent price changes represent random departures from previous prices. The logic of the random walk idea is that if the flow of information is unimpeded and information is immediately reflected in stock prices, then tomorrow’s price change will reflect only tomorrow’s news and will be independent of the price changes today. But news is by definition unpredictable, and, thus, resulting price changes must be unpredictable and random. As a result, prices fully reflect all known information, and even uninformed investors buying a diversified portfolio at the tableau of prices given by the market will obtain a rate of return as generous as that achieved by the experts.

■ Burton G. Malkiel is Chemical Bank Chairman’s Professor of Economics, Princeton University, Princeton, New Jersey. His e-mail address is {bmalkiel@princeton.edu}. 
The way I put it in my book, *A Random Walk Down Wall Street*, first published in 1973, a blindfolded chimpanzee throwing darts at the *Wall Street Journal* could select a portfolio that would do as well as the experts. Of course, the advice was not literally to throw darts, but instead to throw a towel over the stock pages—that is, to buy a broad-based index fund that bought and held all the stocks in the market and that charged very low expenses.

By the start of the twenty-first century, the intellectual dominance of the efficient market hypothesis had become far less universal. Many financial economists and statisticians began to believe that stock prices are at least partially predictable. A new breed of economists emphasized psychological and behavioral elements of stock-price determination, and they came to believe that future stock prices are somewhat predictable on the basis of past stock price patterns as well as certain “fundamental” valuation metrics. Moreover, many of these economists were even making the far more controversial claim that these predictable patterns enable investors to earn excess risk adjusted rates of return.

This paper examines the attacks on the efficient market hypothesis and the belief that stock prices are partially predictable. While I make no attempt to present a complete survey of the purported regularities or anomalies in the stock market, I will describe the major statistical findings as well as their behavioral underpinnings, where relevant, and also examine the relationship between predictability and efficiency. I will also describe the major arguments of those who believe that markets are often irrational by analyzing the “crash of 1987,” the Internet “bubble” of the *fin de siecle* and other specific irrationalities often mentioned by critics of efficiency. I conclude that our stock markets are far more efficient and far less predictable than some recent academic papers would have us believe. Moreover, the evidence is overwhelming that whatever anomalous behavior of stock prices may exist, it does not create a portfolio trading opportunity that enables investors to earn extraordinary risk adjusted returns.

At the outset, it is important to make clear what I mean by the term “efficiency.” I will use as a definition of efficient financial markets that such markets do not allow investors to earn above-average returns without accepting above-average risks. A well-known story tells of a finance professor and a student who come across a $100 bill lying on the ground. As the student stops to pick it up, the professor says, “Don’t bother—if it were really a $100 bill, it wouldn’t be there.” The story well illustrates what financial economists usually mean when they say markets are efficient. Markets can be efficient in this sense even if they sometimes make errors in valuation, as was certainly true during the 1999–early 2000 Internet “bubble.” Markets can be efficient even if many market participants are quite irrational. Markets can be efficient even if stock prices exhibit greater volatility than can apparently be explained by fundamentals such as earnings and dividends. Many of us economists who believe in efficiency do so because we view markets as amazingly successful devices for reflecting new information rapidly and, for the most part, accurately. Above all, we believe that financial markets are efficient because they don’t allow investors to earn above-average risk adjusted returns. In short, we
believe that $100 bills are not lying around for the taking, either by the professional or the amateur investor.

What I do not argue is that the market pricing is always perfect. After the fact, we know that markets have made egregious mistakes, as I think occurred during the recent Internet “bubble.” Nor do I deny that psychological factors influence securities prices. But I am convinced that Benjamin Graham (1965) was correct in suggesting that while the stock market in the short run may be a voting mechanism, in the long run it is a weighing mechanism. True value will win out in the end. Before the fact, there is no way in which investors can reliably exploit any anomalies or patterns that might exist. I am skeptical that any of the “predictable patterns” that have been documented in the literature were ever sufficiently robust so as to have created profitable investment opportunities, and after they have been discovered and publicized, they will certainly not allow investors to earn excess returns.

A Nonrandom Walk Down Wall Street

In this section, I review some of the patterns of possible predictability suggested by studies of the behavior of past stock prices.

Short-Term Momentum, Including Underreaction to New Information

The original empirical work supporting the notion of randomness in stock prices looked at measures of short-run serial correlations between successive stock price changes. In general, this work supported the view that the stock market has no memory—that is, the way a stock price behaved in the past is not useful in divining how it will behave in the future; for example, see the survey of articles contained in Cootner (1964). More recent work by Lo and MacKinlay (1999) finds that short-run serial correlations are not zero and that the existence of “too many” successive moves in the same direction enable them to reject the hypothesis that stock prices behave as true random walks. There does seem to be some momentum in short-run stock prices. Moreover, Lo, Mamaysky and Wang (2000) also find, through the use of sophisticated nonparametric statistical techniques that can recognize patterns, some of the stock price signals used by “technical analysts,” such as “head and shoulders” formations and “double bottoms,” may actually have some modest predictive power.

Economists and psychologists in the field of behavioral finance find such short-run momentum to be consistent with psychological feedback mechanisms. Individuals see a stock price rising and are drawn into the market in a kind of “bandwagon effect.” For example, Shiller (2000) describes the rise in the U.S. stock market during the late 1990s as the result of psychological contagion leading to irrational exuberance. The behavioralists offered another explanation for patterns of short-run momentum—a tendency for investors to underreact to new information. If the full impact of an important news announcement is only grasped over a period of time, stock prices will exhibit the positive serial correlation found by
investigators. As behavioral finance became more prominent as a branch of the study of financial markets, momentum, as opposed to randomness, seemed reasonable to many investigators.

However, several factors should prevent us from interpreting the empirical results reported above as an indication that markets are inefficient. First, while the stock market may not be a mathematically perfect random walk, it is important to distinguish statistical significance from economic significance. The statistical dependencies giving rise to momentum are extremely small and are not likely to permit investors to realize excess returns. Anyone who pays transactions costs is unlikely to fashion a trading strategy based on the kinds of momentum found in these studies that will beat a buy-and-hold strategy. Indeed, Odean (1999) suggests that momentum investors do not realize excess returns. Quite the opposite—a sample of such investors suggests that such traders did far worse than buy-and-hold investors even during a period where there was clear statistical evidence of positive momentum. This is because of the large transactions costs involved in attempting to exploit whatever momentum exists. Similarly, Lesmond, Schill and Zhou (2001) find that standard “relative strength” strategies are not profitable because of the trading costs involved in their execution.

Second, while behavioral hypotheses about bandwagon effects and underreaction to new information may sound plausible enough, the evidence that such effects occur systematically in the stock market is often rather thin. For example, Eugene Fama (1998) surveys the considerable body of empirical work on “event studies” that seeks to determine if stock prices respond efficiently to information. The “events” include such announcements as earnings surprises, stock splits, dividend actions, mergers, new exchange listings and initial public offerings. Fama finds that apparent underreaction to information is about as common as overreaction, and postevent continuation of abnormal returns is as frequent as postevent reversals. He also shows that many of the return “anomalies” arise only in the context of some very particular model and that the results tend to disappear when exposed to different models for expected “normal” returns, different methods to adjust for risk and when different statistical approaches are used to measure them. For example, a study that gives equal weight to postannouncement returns of many stocks can produce different results from a study that weights the stocks according to their value. Certainly, whatever momentum displayed by stock prices does not appear to offer investors a dependable way to earn abnormal returns.

The key factor is whether any patterns of serial correlation are consistent over time. Momentum strategies, which refer to buying stocks that display positive serial correlation and/or positive relative strength, appeared to produce positive relative returns during some periods of the late 1990s, but highly negative relative returns during 2000. It is far from clear that any stock price patterns are useful for investors in fashioning an investment strategy that will dependably earn excess returns.

Many predictable patterns seem to disappear after they are published in the finance literature. Schwert (2001) points out two possible explanations for such a pattern. One explanation may be that researchers are always sifting through
mountains of financial data. Their normal tendency is to focus on results that challenge perceived wisdom, and every now and again, a combination of a certain sample and a certain technique will produce a statistically significant result that seems to challenge the efficient markets hypothesis. Alternatively, perhaps practitioners learn quickly about any true predictable pattern and exploit it to the extent that it becomes no longer profitable. My own view is that such apparent patterns were never sufficiently large or stable to guarantee consistently superior investment results, and certainly, such patterns will never be useful for investors after they have received considerable publicity. The so-called “January effect,” for example, in which stock prices rose in early January, seems to have disappeared soon after it was discovered.

**Long-Run Return Reversals**

In the short-run, when stock returns are measured over periods of days or weeks, the usual argument against market efficiency is that some positive serial correlation exists. But many studies have shown evidence of negative serial correlation—that is, return reversals—over longer holding periods. For example, Fama and French (1988) found that 25 to 40 percent of the variation in long holding period returns can be predicted in terms of a negative correlation with past returns. Similarly, Poterba and Summers (1988) found substantial mean reversion in stock market returns at longer horizons.

Some studies have attributed this forecastability to the tendency of stock market prices to “overreact.” DeBondt and Thaler (1985), for example, argue that investors are subject to waves of optimism and pessimism that cause prices to deviate systematically from their fundamental values and later to exhibit mean reversion. They suggest that such overreaction to past events is consistent with the behavioral decision theory of Kahneman and Tversky (1979), where investors are systematically overconfident in their ability to forecast either future stock prices or future corporate earnings. These findings give some support to investment techniques that rest on a “contrarian” strategy, that is, buying the stocks, or groups of stocks, that have been out of favor for long periods of time and avoiding those stocks that have had large run-ups over the last several years.

There is indeed considerable support for long-run negative serial correlation in stock returns. However, the finding of mean reversion is not uniform across studies and is quite a bit weaker in some periods than it is for other periods. Indeed, the strongest empirical results come from periods including the Great Depression—which may be a time with patterns that do not generalize well. Moreover, such return reversals for the market as a whole may be quite consistent with the efficient functioning of the market since they could result, in part, from the volatility of interest rates and the tendency of interest rates to be mean reverting. Since stock returns must rise or fall to be competitive with bond returns, there is a tendency when interest rates go up for prices of both bond and stocks to go down, and as interest rates go down for prices of bonds and stocks to go up. If interest rates revert to the mean over time, this pattern will tend to generate return
reversals, or mean reversion, in a way that is quite consistent with the efficient functioning of markets.

Moreover, it may not be possible to profit from the tendency for individual stocks to exhibit return reversals. Fluck, Malkiel and Quandt (1997) simulated a strategy of buying stocks over a 13-year period during the 1980s and early 1990s that had particularly poor returns over the past three to five years. They found that stocks with very low returns over the past three to five years had higher returns in the next period and that stocks with very high returns over the past three to five years had lower returns in the next period. Thus, they confirmed the very strong statistical evidence of return reversals. However, they also found that returns in the next period were similar for both groups, so they could not confirm that a contrarian approach would yield higher-than-average returns. There was a statistically strong pattern of return reversal, but not one that implied an inefficiency in the market that would enable investors to make excess returns.

Seasonal and Day-of-the-Week Patterns

A number of researchers have found that January has been a very unusual month for stock market returns. Returns from an equally weighted stock index have tended to be unusually high during the first two weeks of the year. The return premium has been particularly evident for stocks with relatively small total capitalizations (Keim, 1983). Haugen and Lakonishok (1988) documented the high January returns in a book titled *The Incredible January Effect*. There also appear to be a number of day-of-the-week effects. For example, French (1980) documents significantly higher Monday returns. There appear to be significant differences in average daily returns in countries other than the United States (Hawawini and Keim, 1995). There also appear to be some patterns in returns around the turn of the month (Lakonishok and Smidt, 1988), as well as around holidays (Ariel, 1990).

The general problem with these predictable patterns or anomalies, however, is that they are not dependable from period to period. Wall Street traders now joke that the “January effect” is more likely to occur on the previous Thanksgiving. Moreover, these nonrandom effects (even if they were dependable) are very small relative to the transactions costs involved in trying to exploit them. They do not appear to offer arbitrage opportunities that would enable investors to make excess risk adjusted returns.

Predictable Patterns Based on Valuation Parameters

Considerable empirical research has been conducted to determine if future stock returns can be predicted on the basis of initial valuation parameters. It is claimed that valuation ratios, such as the price-earnings multiple or the dividend yield of the stock market as a whole, have considerable predictive power. This section examines the body of work based on time series analyses.
Predicting Future Returns from Initial Dividend Yields

Formal statistical tests of the ability of dividend yields (that is, the ratio of dividend to stock price) to forecast future returns have been conducted by Fama and French (1988) and Campbell and Shiller (1988). Depending on the forecast horizon involved, as much as 40 percent of the variance of future returns for the stock market as a whole can be predicted on the basis of the initial dividend yield of the market index.

An interesting way of presenting the results is shown in the top panel of Exhibit 1. The exhibit was produced by measuring the dividend yield of the broad U.S. stock market Standard & Poor’s 500 Stock Index each quarter since 1926 and then calculating the market’s subsequent ten-year total return through the year 2001. The observations were then divided into deciles depending upon the level of the initial dividend yield. In general, the exhibit shows that investors have earned a higher rate of return from the stock market when they purchased a market basket of equities with an initial dividend yield that was relatively high and relatively low future rates of return when stocks were purchased at low dividend yields.

These findings are not necessarily inconsistent with efficiency. Dividend yields of stocks tend to be high when interest rates are high, and they tend to be low when interest rates are low. Consequently, the ability of initial yields to predict returns may simply reflect the adjustment of the stock market to general economic conditions. Moreover, the use of dividend yields to predict future returns has been ineffective since the mid-1980s. Dividend yields have been at the 3 percent level or below continuously since the mid-1980s, indicating very low forecasted returns. In fact, for all ten-year periods from 1985 through 1992 that ended June 30, 2002, realized annual equity returns from the market index have averaged over 15 percent. One possible explanation is that the dividend behavior of U.S. corporations may have changed over time (Bagwell and Shoven, 1989; Fama and French, 2001). Companies in the twenty-first century may be more likely to institute a share repurchase program rather than increase their dividends. Thus, dividend yield may not be as meaningful as in the past as a useful predictor of future equity returns.

Finally, it is worth noting that this phenomenon does not work consistently with individual stocks (Fluck, Malkiel and Quandt, 1997). Investors who simply purchase a portfolio of individual stocks with the highest dividend yields in the market will not earn a particularly high rate of return. One popular implementation of such a “high dividend” strategy in the United States is the “Dogs of the Dow Strategy,” which involves buying the ten stocks in the Dow Jones Industrial Average with the highest dividend yields. For some past periods, this strategy handily outpaced the overall average, and so several “Dogs of the Dow” mutual funds were brought to market and aggressively sold to individual investors. However, such funds generally underperformed the market averages during the 1995–1999 period.

Predicting Market Returns from Initial Price-Earnings Multiples

The same kind of predictability for the market as a whole, as was demonstrated for dividends, has been shown for price-earnings ratios. The data are shown in the
Exhibit 1
The Future 10-Year Rates of Return When Stocks are Purchased at Alternative Initial Dividend Yields (D/P)

The Future 10-Year Rates of Return When Stocks are Purchased at Alternative Initial Price-to-Earnings (P/E) Multiples

Source: The Leuthold Group

bottom half of Exhibit 1. The exhibit presents a decile analysis similar to that described for dividend yields above. Investors have tended to earn larger long-horizon returns when purchasing the market basket of stocks at relatively low price-earnings multiples. Campbell and Shiller (1998) report that initial P/E ratios
explained as much as 40 percent of the variance of future returns. They conclude that equity returns have been predictable in the past to a considerable extent.

Consider, however, the recent experience of investors who have attempted to undertake investment strategies based either on the level of the price-earnings multiple or the dividend yield to predict future long-horizon returns. Price-earnings multiples for the Standard & Poor’s 500 stock index rose into the low 20s on June 30, 1987 (suggesting very low long-horizon returns). Dividend yields fell below 3 percent. Price-earnings multiples rose into the low 20s. The average annual total return from the index over the next 10 years was an extraordinarily generous 16.7 percent. Dividend yields, again, fell to 3 percent in June 1992. Price-earnings multiples rose to the mid-20s. The subsequent return through June 2002 was 11.4 percent. The yield of the index fluctuated between 2 and 3 percent from 1993 through 1995 and earnings multiples remained in the mid-20s, yet long-horizon returns through June 30, 2002, fluctuated between 11 and 12 percent. Even from early December 1996, the date of Campbell and Shiller’s presentation to the Federal Reserve suggesting near zero returns for the S&P500, the index provided almost a 7 percent annual return through mid-2002. Such results suggest to me a very cautious assessment of the extent to which stock market returns are predictable on this basis.

Other Predictable Time Series Patterns

Studies have found some amount of predictability of stock returns based on various financial statistics. For example, Fama and Schwert (1977) found that short-term interest rates were related to future stock returns. Campbell (1987) found that term structure of interest rates spreads contained useful information for forecasting stock returns. Keim and Stambaugh (1986) found that risk spreads between high-yield corporate bonds and short rates had some predictive power. Again, even if some predictability exists, it may reflect time varying risk premiums and required rates of return for stock investors rather than an inefficiency. Moreover, it is far from clear that any of these results can be used to generate profitable trading strategies. Whether such historical statistical relations give investors reliable and useful guides to appropriate asset allocation is far from clear.

Cross-Sectional Predictable Patterns Based on Firm Characteristics and Valuation Parameters

A large number of patterns that are claimed to be predictable are based on firm characteristics and different valuation parameters.

The Size Effect

One of the strongest effects investigators have found is the tendency over long periods of time for smaller-company stocks to generate larger returns that those of
large-company stocks. Since 1926, small-company stocks in the United States have produced annual rates of return over 1 percentage point larger than the returns from large stocks (Keim, 1983). Fama and French (1993) examined data from 1963 to 1990 and divided all stocks into deciles according to their size as measured by total capitalization. Decile one contained the smallest 10 percent of all stocks, while decile ten contained the largest stocks. The results, plotted in Exhibit 2, show a clear tendency for the deciles made up of portfolios of smaller stocks to generate higher average monthly returns than deciles made up of larger stocks.

The crucial issue here is the extent to which the higher returns of small companies represent a predictable pattern that will allow investors to generate excess risk-adjusted returns. According to the capital asset pricing model, the correct measure of risk for a stock is its “beta”—that is, the extent to which the return of the stock varies with the return for the market as a whole. If the beta measure of systematic risk from the capital asset pricing model is accepted as the correct risk measurement statistic, the size effect can be interpreted as indicating an anomaly and a market inefficiency, because using this measure portfolios consisting of smaller stocks have excess risk-adjusted returns. Fama and French (1993) point out, however, that the average relationship between beta and return during the 1963–1990 period was flat—not upward sloping as the capital asset pricing model predicts. Moreover, if stocks are divided up by beta deciles, ten portfolios constructed by size display the same kind of positive relationship shown in Exhibit 2. On the other hand, within size deciles, the relationship between beta and return continues to be flat. Fama and French suggest that size may be a far better proxy for risk than beta, and therefore that their findings should not be interpreted as indicating that markets are inefficient.

The dependability of the size phenomenon is also open to question. From the mid-1980s through the decade of the 1990s, there has been no gain from holding smaller stocks. Indeed, in most world markets, larger capitalization stocks produced larger rates of return. It may be that the growing institutionalization of the market led portfolio managers to prefer larger companies with more liquidity to smaller companies where it would be difficult to liquidate significant blocks of stock. Finally, it is also possible that some studies of the small-firm effect have been affected by survivorship bias. Today’s computerized databases of companies include only small firms that have survived, not the ones that later went bankrupt. Thus, a researcher who examined the ten-year performance of today’s small companies would be measuring the performance of those companies that survived—not the ones that failed.

**Value Stocks**

Several studies suggest that “value” stocks have higher returns than so-called “growth” stocks. The most common two methods of identifying value stocks have been price-earnings ratios and price-to-book-value ratios.

Stocks with low price-earnings multiples (often called “value” stocks) appear to provide higher rates of return than stocks with high price-to-earnings ratios, as first
shown by Nicholson (1960) and later confirmed by Ball (1978) and Basu (1983). This finding is consistent with the views of behavioralists that investors tend to be overconfident of their ability to project high earnings growth and thus overpay for “growth” stocks (for example, Kahneman and Riepe, 1998). The finding is also consistent with the views of Graham and Dodd (1934), first expounded in their classic book on security analysis and later championed by the legendary U.S. investor Warren Buffett. Similar results have been shown for price/cash flow multiples, where cash flow is defined as earnings plus depreciation and amortization (Hawawini and Keim, 1995).

The ratio of stock price to book value, defined as the value of a firm’s assets minus its liabilities divided by the number of shares outstanding, has also been found to be a useful predictor of future returns. Low price-to-book is considered to be another hallmark of so-called “value” in equity securities and is also consistent with the view of behavioralists that investors tend to overpay for “growth” stocks that subsequently fail to live up to expectations. Fama and French (1993) concluded that size and price-to-book-value together provide considerable explanatory power for future returns, and once they are accounted for, little additional influence can be attributed to P/E multiples. Fama and French (1997) also conclude that the P/BV effect is important in many world stock markets other than the United States.

Such results raise questions about the efficiency of the market if one accepts the capital asset pricing model, as Lakonishok, Shleifer and Vishny (1994) point out. But these findings do not necessarily imply inefficiency. They may simply indicate failure of the capital asset pricing model to capture all the dimensions of

---

**Exhibit 2**

**Average Monthly Returns for Portfolios Formed on the Basis of Size: 1963–1990**

<table>
<thead>
<tr>
<th>Size</th>
<th>Small Stocks</th>
<th>Large Stocks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.8</td>
<td>.9</td>
</tr>
<tr>
<td>2</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>3</td>
<td>1.1</td>
<td>1.1</td>
</tr>
<tr>
<td>4</td>
<td>1.2</td>
<td>1.2</td>
</tr>
<tr>
<td>5</td>
<td>1.3</td>
<td>1.3</td>
</tr>
<tr>
<td>6</td>
<td>1.4</td>
<td>1.4</td>
</tr>
<tr>
<td>7</td>
<td>1.5</td>
<td>1.5</td>
</tr>
<tr>
<td>8</td>
<td>1.6</td>
<td>1.6</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Source: Fama and French (1992).*
risk. For example, Fama and French (1993) suggest that the price-to-book-value ratio may reflect another risk factor that is priced into the market and not captured by the capital asset pricing model. Companies in some degree of financial distress, for example, are likely to sell at low prices relative to book values. Fama and French (1993) argue that a three-factor asset-pricing model (including price-to-book-value and size as measures of risk) is the appropriate benchmark against which anomalies should be measured.

We also need to keep in mind that the results of published studies—even those done over decades—may still be time-dependent and ask whether the return patterns of academic studies can actually be generated with real money. Exhibit 3 presents average actual returns generated by mutual funds classified by either their “growth” or “value” objectives. “Value” funds are so classified if they buy stocks with price-to-earnings and price-to-book-value multiples that are below the averages for the whole stock market. Over a period running back to the 1930s, it does not appear that investors could actually have realized higher rates of return from mutual funds specializing in “value” stocks. Indeed, the exhibit suggests that the Fama-French (1993) period from the early 1960s through 1990 may have been a unique period in which value stocks rather consistently produced higher rates of return.

Schwert (2001) points out that the investment firm of Dimensional Fund Advisors actually began a mutual fund that selected value stocks quantitatively according to the Fama and French (1993) criteria. The abnormal return of such a portfolio (adjusting for beta, the capital asset pricing model measure of risk) was a negative 0.2 percent per month over the 1993–1998 period. The absence during that period of an excess return to the “value” stocks is consistent with the results from “actively managed” value mutual funds shown in Exhibit 3.

The Equity Risk Premium Puzzle

Another puzzle that is often used to suggest that markets are less than fully rational is the existence of a very large historical equity risk premium that seems inconsistent with the actual riskiness of common stocks as can be measured statistically. For example, using the Ibbotson data on stock returns from 1926 through 2001, common stocks have produced rates of return of approximately 10.5 percent, while high-grade bonds have returned only about 5.5 percent. I believe that this finding is the result of a combination of perceived equity risk being considerably higher during the early years of the period and of average equity returns being much higher than had been forecast by investors.

It is easy to say 50 to 75 years later that common stocks were underpriced during the 1930s and 1940s. But remember that the annual average of almost 6 percent growth in corporate earnings and dividends that has occurred since 1926 was hardly a foregone conclusion during a period of severe depression and world war. Indeed, the U.S. stock market is almost unique in that it is one of the few world markets that remained in continuous operation during the entire period and the measured risk premium results, in part, from survivorship bias. One must be very
careful to distinguish between expected risk premiums and such premiums measured after the fact. Fama and French (2002) argue that the high average realized returns result in part from large unexpected capital gains. Economists such as Shiller have suggested that during the early 2000s, the expected equity risk premium was, if anything, irrationally too low.

**Exhibit 3**

Reversion to the Mean: Relative Performance of “Value” vs. “Growth” Mutual Funds, 1937–June 2002

Source: Lipper Analytic Services and Bogle Research Institute Valley Forge, Pennsylvania.  
*Note:* The exhibit shows the cumulative value of one dollar invested in the average “value” fund divided by the same statistic calculated for the average “growth” fund.

Summarizing the “Anomalies” and Predictable Patterns

The preceding sections have pointed out many “anomalies” and statistically significant predictable patterns in the stock returns that have been uncovered in the literature. However, these patterns are not robust and dependable in different sample periods, and some of the patterns based on fundamental valuation measures of individual stocks may simply reflect better proxies for measuring risk.

Moreover, many of these patterns, even if they did exist, could self-destruct in the future, as many of them have already done. Indeed, this is the logical reason why one should be cautious not to overemphasize these anomalies and predictable patterns. Suppose, for example, one of the anomalies or predictable patterns appears to be robust. Suppose there is a truly dependable and exploitable January effect, that the stock market—especially stocks of small companies—will generate extraordinary returns during the first five days of January. What will investors do? They will buy on the last day of December and sell on January 5. But then investors find that the market rallied on the last day of December, and so they will need to begin to buy on the next-to-last day of December; and because there is so much “profit taking” on January 5, investors will have to sell on January 4 to take advantage of this effect. Thus, to beat the gun, investors will have to be buying
earlier and earlier in December and selling earlier and earlier in January so that eventually the pattern will self-destruct. Any truly repetitive and exploitable pattern that can be discovered in the stock market and can be arbitrated away will self-destruct. Indeed, the January effect became undependable after it received considerable publicity.

Similarly, suppose there is a general tendency for stock prices to underreact to certain new events, leading to abnormal returns to investors who exploit the lack of full immediate adjustment (DeBondt and Thaler, 1995; Campbell, Lo and MacKinlay, 1977). “Quantitative” investment managers will then develop trading strategies to exploit the pattern. Indeed, the more potentially profitable a discoverable pattern is, the less likely it is to survive.

Many of the predictable patterns that have been discovered may simply be the result of data mining. The ease of experimenting with financial databanks of almost every conceivable dimension makes it quite likely that investigators will find some seemingly significant but wholly spurious correlation between financial variables or among financial and nonfinancial data sets. Given enough time and massaging of data series, it is possible to tease almost any pattern out of most data sets. Moreover, the published literature is likely to be biased in favor of reporting such results. Significant effects are likely to be published in professional journals while negative results, or boring confirmations of previous findings, are relegated to the file drawer or discarded. Data-mining problems are unique to nonexperimental sciences, such as economics, which rely on statistical analysis for their insights and cannot test hypotheses by running repeated controlled experiments.

An exchange at a symposium about a decade ago between Robert Shiller, an economist who is sympathetic to the argument that stock prices are partially predictable and skeptical about market efficiency, and Richard Roll, an academic financial economist who also is a portfolio manager, is quite revealing (Roll and Shiller, 1992). After Shiller stressed the importance of inefficiencies in the pricing of stocks, Roll responded as follows:

I have personally tried to invest money, my client’s money and my own, in every single anomaly and predictive device that academics have dreamed up. . . . I have attempted to exploit the so-called year-end anomalies and a whole variety of strategies supposedly documented by academic research. And I have yet to make a nickel on any of these supposed market inefficiencies . . . a true market inefficiency ought to be an exploitable opportunity. If there’s nothing investors can exploit in a systematic way, time in and time out, then it’s very hard to say that information is not being properly incorporated into stock prices.

**Seemingly Irrefutable Cases of Inefficiency**

Critics of efficiency argue that there are several instances of recent market history where market prices could not plausibly have been set by rational investors
and that psychological considerations must have played the dominant role. It is alleged, for example, that the stock market lost about one-third of its value from early to mid-October 1987 with essentially no change in the general economic environment. How could market prices be efficient both at the start of October and during the middle of the month? Similarly, it is widely believed that the pricing of Internet stocks in early 2000 could only be explained by the behavior of irrational investors. Do such events make a belief in efficient markets untenable?

The Market Crash of October 1987

Can the October 1987 market crash be explained by rational considerations, or does such a rapid and significant change in market valuations prove the dominance of psychological rather than logical factors in understanding the stock market? Behaviorists would say that the one-third drop in market prices, which occurred early in October 1987, can only be explained by relying on psychological considerations, since the basic elements of the valuation equation did not change rapidly over that period. It is, of course, impossible to rule out the existence of behavioral or psychological influences on stock market pricing. But logical considerations can explain a sharp change in market valuations such as occurred during the first weeks of October 1987.

A number of factors could rationally have changed investors’ views about the proper value of the stock market in October 1987. For one thing, yields on long-term Treasury bonds increased from about 9 percent to almost 10.5 percent in the two months prior to mid-October. Moreover, a number of events may rationally have increased risk perceptions during the first two weeks of October. Early in the month, Congress threatened to impose a “merger tax” that would have made merger activity prohibitively expensive and could well have ended the merger boom. The risk that merger activity might be curtailed increased risks throughout the stock market by weakening the discipline over corporate management that potential takeovers provide. Also, in early October 1987, then Secretary of the Treasury James Baker had threatened to encourage a further fall in the exchange value of the dollar, increasing risks for foreign investors and frightening domestic investors as well. While it is impossible to correlate each day’s movement in stock prices to specific news events, it is not unreasonable to ascribe the sharp decline in mid-October to the cumulative effect of a number of unfavorable “fundamental” events. As Merton Miller (1991) has written, “. . . on October 19, some weeks of external events, minor in themselves. . . . cumulatively signaled a possible change in what had been up to then a very favorable political and economic climate for equities. . . . and. . . . many investors simultaneously came to believe they were holding too large a share of their wealth in risky equities.”

Share prices can be highly sensitive as a result of rational responses to small changes in interest rates and risk perceptions. Suppose stocks are priced as the present value of the expected future stream of dividends. For a long-term holder of stocks, this rational principle of valuation translates to a formula:
\[ r = D/P + g, \]

where \( r \) is the rate of return, \( D/P \) is the (expected) dividend yield, and \( g \) is the long-term growth rate. For present purposes, consider \( r \) to be the required rate of return for the market as a whole. Suppose initially that the “riskless” rate of interest on government bonds is 9 percent and that the required additional risk premium for equity investors is 2 percentage points. In this case, \( r \) will be 11 percent \((0.09 + 0.02 = 0.11)\). If a typical stock’s expected growth rate, \( g \), is 7 percent and if the dividend is $4 per share, we can solve for the appropriate price of the stock index \((P)\), obtaining

\[ 0.11 = \frac{4}{P} + 0.07 \]

\[ P = $100. \]

Now assume that yields on government bonds rise from 9 to 10.5 percent, with no increase in expected inflation, and that risk perceptions increase so that stock-market investors now demand a premium of 2.5 percentage points instead of the 2 points in the previous example. The appropriate rate of return or discount rate for stocks, \( r \), rises then from 11 percent to 13 percent \((0.105 + 0.025)\), and the price of our stock index falls from $100 to $66.67:

\[ 0.13 = \frac{4}{P} + 0.07 \]

\[ P = $66.67. \]

The price must fall to raise the dividend yield from 4 to 6 percent so as to raise the total return by the required 2 percentage points. Clearly, no irrationality is required for share prices to suffer quite dramatic declines with the sorts of changes in interest rates and risk perceptions that occurred in October 1987. Of course, even a very small decline in anticipated growth would have magnified these declines in warranted share valuations.

This is not to say that psychological factors were irrelevant in explaining the sharp drop in prices during October 1987—they undoubtedly played a role. But it would be a mistake to dismiss the significant change in the external environment, which can provide an entirely rational explanation for a significant decline in the appropriate values for common stocks.

**The Internet Bubble of the Late 1990s**

Another stock market event often cited by behavioralists as clear evidence of the irrationality of markets is the Internet “bubble” of the late 1990s. Surely, the
remarkable market values assigned to Internet and related high-tech companies seem inconsistent with rational valuation. I have some sympathy with behavioralists in this instance, and in reviewing Robert Shiller’s (2000) *Irrational Exuberance*, I agreed that it was in the high-tech sector of the market that his thesis could be supported. But even here, when we know after the fact that major errors were made, there were certainly no arbitrage opportunities available to rational investors before the “bubble” popped.

Equity valuations rest on uncertain future forecasts. Even if all market participants rationally price common stocks as the present value of all future cash flows expected, it is still possible for excesses to develop. We know now, with the benefit of hindsight, that outlandish and unsupportable claims were being made regarding the growth of the Internet (and the related telecommunications structure needed to support it). We know now that projections for the rates and duration of growth of these “new economy” companies were unsustainable. But remember, sharp-penciled professional investors argued that the valuations of high-tech companies were proper. Many of Wall Street’s most respected security analysts, including those independent of investment banking firms, were recommending Internet stocks to the firm’s institutional and individual clients as being fairly valued. Professional pension fund and mutual fund managers overweighted their portfolios with high-tech stocks.

While it is now clear in retrospect that such professionals were egregiously wrong, there was certainly no obvious arbitrage opportunity available. One could disagree with the projected growth rates of security analysts. But who could be sure, with the use of the Internet for a time doubling every several months, that the extraordinary growth rates that could justify stock valuations were impossible? After all, even Alan Greenspan was singing the praises of the new economy. Nothing is ever as clear in prospect as it is in retrospect. The extent of the “bubble” was only clear in retrospect.

Not only is it almost impossible to judge with confidence what the proper fundamental value is for any security, but potential arbitrageurs face additional risks. Shleifer (2000) has argued that “noise trader risk”—the risk from traders who are attempting to buy into rising markets and sell into declining markets—limits the extent to which one should expect arbitrage to bring prices quickly back to rational values even in the presence of an apparent bubble. Professional arbitrageurs will be loath to sell short a stock they believe is trading at two times its “fundamental” value when it is always possible that some greater fools may be willing to pay three times the stock’s value. Arbitrageurs are quite likely to have short horizons, since even temporary losses may induce their clients to withdraw their money.

While there were no profitable and predictable arbitrage opportunities available during the Internet “bubble,” and while stock prices eventually did adjust to levels that more reasonably reflected the likely present value of their cash flows, an argument can be maintained that the asset prices did remain “incorrect” for a period of time. The result was that too much new capital flowed to Internet and
related telecommunications companies. Thus, the stock market may well have temporarily failed in its role as an efficient allocator of equity capital. Fortunately, “bubble” periods are the exception rather than the rule, and acceptance of such occasional mistakes is the necessary price of a flexible market system that usually does a very effective job of allocating capital to its most productive uses.

Other Illustrations of Irrational Pricing

Are there not some illustrations of irrational pricing that can be clearly ascertained as they arise, not simply after a “bubble” has burst? My favorite illustration concerns the spinoff of Palm Pilot from its parent 3Com Corporation during the height of the Internet boom in early 2000. Initially, only 5 percent of the Palm Pilot shares were distributed to the public; the other 95 percent remained on 3Com’s balance sheet. As Palm Pilot began trading, enthusiasm for the shares was so great that the 95 percent of its shares still owned by 3Com had a market value considerably more than the entire market capitalization of 3Com, implying that all the rest of its business had a negative value. Other illustrations involve ticker symbol confusion. Rasches (2001) finds clear evidence of comovement of stocks with similar ticker symbols; for example, the stock of MCI Corporation (ticker symbol MCIC) moves in tandem with an unrelated closed-end bond investment fund Mass Mutual Corporate Investors (ticker symbol MCI). In a charming article entitled “A Rose.com by Any Other Name,” Cooper, Dimitrov and Rau (2001) found positive stock price reactions during 1998 and 1999 on corporate name changes when “dot com” was added to the corporate title. Finally, it has been argued that closed-end funds sell at irrational discounts from their net asset values (for example, Shleifer, 2000).

But none of these illustrations should shake our faith that exploitable arbitrage opportunities should not exist in an efficient market. The apparent arbitrage in the Palm Pilot case (sell Palm Pilot short and buy 3Com) could not be undertaken because not enough Palm stock was outstanding to make borrowing the stock possible to effectuate a short sale. The “anomaly” disappeared once 3Com spun off more of Palm stock. Moreover, the potential profits from name or ticker symbol confusion are extremely small relative to the transactions costs that would be required to exploit them. Finally, the “closed-end fund puzzle” is not really a puzzle today. Discounts have narrowed from historical averages for funds with assets traded in liquid markets, and researchers such as Ross (forthcoming) have suggested that they can largely be explained by fund management fees. Perhaps the more important puzzle today is why so many investors buy high-expense, actively managed mutual funds instead of low-cost index funds.

The Performance of Professional Investors

For me, the most direct and most convincing tests of market efficiency are direct tests of the ability of professional fund managers to outperform the market
as a whole. Surely, if market prices were determined by irrational investors and systematically deviated from rational estimates of the present value of corporations, and if it were easy to spot predictable patterns in security returns or anomalous security prices, then professional fund managers should be able to beat the market. Direct tests of the actual performance of professionals, who often are compensated with strong incentives to outperform the market, should represent the most compelling evidence of market efficiency.

A remarkably large body of evidence suggests that professional investment managers are not able to outperform index funds that buy and hold the broad stock market portfolio. The first study of mutual fund performance was undertaken by Jensen (1968). He found that active mutual fund managers were unable to add value and, in fact, tended to underperform the market by approximately the amount of their added expenses. I repeated Jensen’s study with data from a subsequent period and confirmed the earlier results (Malkiel, 1995). Moreover, I found that the degree of “survivorship bias” in the data was substantial; that is, poorly performing funds tend to be merged into other funds in the mutual fund’s family complex, thus burying the records of many of the underperformers. Exhibit 4 updates the study I performed through mid-2002, showing that the return for surviving funds is quite a bit better than the actual return for all funds, including funds liquidated or merged out of existence. Survivorship bias makes the interpretation of long-run mutual fund data sets very difficult. But even using data sets with some degree of survivorship bias, one cannot sustain the argument that professional investors can beat the market.

Exhibit 5 presents the percentage of actively managed mutual funds that have been outperformed by the Standard & Poor’s 500 and the Wilshire stock indexes. Throughout the past decade, about three-quarters of actively managed funds have
failed to beat the index. Similar results obtain for earlier decades. Exhibit 6 shows that the median large-cap equity fund has underperformed the S&P 500 index by almost 2 percentage points over the past 10-, 15- and 20-year periods. Exhibit 7 shows similar results in different markets and against different benchmarks.

Managed funds are regularly outperformed by broad index funds, with equivalent risk. Moreover, those funds that produce excess returns in one period are not likely to do so in the next. There is no dependable persistence in performance. During the 1970s, the top 20 mutual funds enjoyed almost double the performance of the index. During the 1980s, those same funds underperformed the index. The best performing funds of the 1980s similarly underperformed during the 1990s. A more dramatic example of the lack of persistence in performance is shown in Exhibit 8. These mutual funds during 1998 and 1999 enjoyed three times the performance of the index. During 2000 and 2001, they did three times worse than the index. Over the long run, the results are even more devastating to active managers. One can count on the fingers of one hand the number of professional portfolio managers who have managed to beat the market by any significant amount. Exhibit 9 shows the distribution of return. Of the original 355 funds, only five of them outperformed the market by 2 percent per year or more.

The record of professionals does not suggest that sufficient predictability exists in the stock market or that there are recognizable and exploitable irrationalities sufficient to produce excess returns.
Exhibit 7

Percentage of Various Actively Managed Funds Outperformed by Benchmark Index 10 Years to 12/31/01

Source: Lipper Analytic Services and Micropal.

Exhibit 8

Getting Burned by Hot Funds

<table>
<thead>
<tr>
<th>Fund Name</th>
<th>Rank</th>
<th>Average Annual Return</th>
<th>Rank</th>
<th>Average Annual Return</th>
</tr>
</thead>
<tbody>
<tr>
<td>Van Wagoner: Emrg Growth</td>
<td>1</td>
<td>105.52</td>
<td>1106</td>
<td>-43.54</td>
</tr>
<tr>
<td>Rydex:OTC Fund; Inv</td>
<td>2</td>
<td>93.43</td>
<td>1103</td>
<td>-36.31</td>
</tr>
<tr>
<td>TCGW Galileo: AGr Eq; Inst</td>
<td>3</td>
<td>92.78</td>
<td>1098</td>
<td>-34.00</td>
</tr>
<tr>
<td>RS Inv: Emerg Growth</td>
<td>4</td>
<td>90.19</td>
<td>1055</td>
<td>-26.17</td>
</tr>
<tr>
<td>PBHG: Large Cap 20</td>
<td>5</td>
<td>84.56</td>
<td>1078</td>
<td>-29.03</td>
</tr>
<tr>
<td>Janus Olympus Fund</td>
<td>6</td>
<td>77.24</td>
<td>1061</td>
<td>-27.03</td>
</tr>
<tr>
<td>Van Kampen Aggr Gro; A</td>
<td>7</td>
<td>76.70</td>
<td>1067</td>
<td>-28.04</td>
</tr>
<tr>
<td>Janus Mercury</td>
<td>8</td>
<td>76.31</td>
<td>1057</td>
<td>-26.35</td>
</tr>
<tr>
<td>PBHG: Sel Equity</td>
<td>9</td>
<td>76.21</td>
<td>1097</td>
<td>-33.19</td>
</tr>
<tr>
<td>WM: Growth; A</td>
<td>10</td>
<td>74.77</td>
<td>1046</td>
<td>-25.82</td>
</tr>
<tr>
<td>Berger New Generatn; Inv</td>
<td>11</td>
<td>73.31</td>
<td>1107</td>
<td>-45.96</td>
</tr>
<tr>
<td>Janus Enterprise</td>
<td>12</td>
<td>72.28</td>
<td>1101</td>
<td>-35.40</td>
</tr>
<tr>
<td>Janus Venture</td>
<td>13</td>
<td>72.22</td>
<td>1091</td>
<td>-30.89</td>
</tr>
<tr>
<td>Fidelity Aggr Growth</td>
<td>14</td>
<td>70.56</td>
<td>1105</td>
<td>-38.02</td>
</tr>
<tr>
<td>Janus Twenty</td>
<td>15</td>
<td>69.90</td>
<td>1090</td>
<td>-30.83</td>
</tr>
<tr>
<td>Amer Cent: New Oppy.</td>
<td>16</td>
<td>67.64</td>
<td>1033</td>
<td>-24.11</td>
</tr>
<tr>
<td>Morg Stan Sm Cap Gro; B</td>
<td>17</td>
<td>66.59</td>
<td>1102</td>
<td>-35.96</td>
</tr>
<tr>
<td>Van Kampen Emerg Gro; A</td>
<td>18</td>
<td>65.67</td>
<td>1021</td>
<td>-22.70</td>
</tr>
<tr>
<td>TCW Galileo: SC Gro; Instl.</td>
<td>19</td>
<td>64.87</td>
<td>1099</td>
<td>-34.77</td>
</tr>
<tr>
<td>BlackRock: MdCp Gro; Instl.</td>
<td>20</td>
<td>64.44</td>
<td>1009</td>
<td>-22.18</td>
</tr>
<tr>
<td>Average Fund Return</td>
<td></td>
<td>76.72</td>
<td></td>
<td>-31.52</td>
</tr>
<tr>
<td>S&amp;P 500 Return</td>
<td></td>
<td>24.75</td>
<td></td>
<td>-10.50</td>
</tr>
</tbody>
</table>

Source: Lipper Analytic Services and Bogle Research Institute, Valley Forge, PA.
Conclusion

As long as stock markets exist, the collective judgment of investors will sometimes make mistakes. Undoubtedly, some market participants are demonstrably less than rational. As a result, pricing irregularities and even predictable patterns in stock returns can appear over time and even persist for short periods. Moreover, the market cannot be perfectly efficient, or there would be no incentive for professionals to uncover the information that gets so quickly reflected in market prices, a point stressed by Grossman and Stiglitz (1980). Undoubtedly, with the passage of time and with the increasing sophistication of our databases and empirical techniques, we will document further apparent departures from efficiency and further patterns in the development of stock returns.

But I suspect that the end result will not be an abandonment of the belief of many in the profession that the stock market is remarkably efficient in its utilization of information. Periods such as 1999 where “bubbles” seem to have existed, at least in certain sectors of the market, are fortunately the exception rather than the rule. Moreover, whatever patterns or irrationalities in the pricing of individual stocks that have been discovered in a search of historical experience are unlikely to persist and will not provide investors with a method to obtain extraordinary returns. If any $100 bills are lying around the stock exchanges of the world, they will not be there for long.

I wish to thank J. Bradford De Long, Timothy Taylor and Michael Waldman for their extremely helpful observations. While they may not agree with all of the conclusions in this paper, they have strengthened my arguments in important ways.
References


