

Discussion of Value Investing:  
The Use of Historical Financial Statement  
Information to Separate Winners from Losers

Wayne Guay  
The Wharton School  
University of Pennsylvania  
2400 Steinberg-Dietrich Hall  
Philadelphia, PA 19104-6365  
email: [guay@wharton.upenn.edu](mailto:guay@wharton.upenn.edu)  
phone: (215) 898-7775  
fax: (215) 573-2054

September, 2000

I would like to thank Phil Berger, John Core, SP Kothari, and Cathy Schrand for their helpful comments.

## *1. Introduction*

Numerous academic studies (e.g., Fama and French [1992], Lakonishok, Shleifer, and Vishny [1994], among others) examine the stock returns of high book-to-market firms ("value" stocks) relative to low book-to-market firms ("glamour" stocks). An empirical regularity of these data is that the returns of value stocks over the last 30 years are significantly greater than the returns of glamour stocks. The reason for this difference has been the topic of much debate. Fama and French [1992] propose that the book-to-market ratio captures a priced element of systematic risk, and that the observed differences in returns between value and glamour stocks reflect a fair compensation for risk. Lakonishok, Shleifer, and Vishny [1994] advance an alternative possibility that high book-to-market firms' stock prices are temporarily depressed because investors over-react to prior poor performance, and maintain expectations about future performance that are "too pessimistic". Fama [1998], as well some conference participants, raise a third possibility that the observed return differential during the past 30 years is a time-period specific regularity that has been uncovered by researchers but has no systematic underlying explanation.

Piotroski [2000] further investigates value stocks and examines whether a simple, accounting-based fundamental analysis strategy, when applied to historical data, can further enhance the returns to investing in high book-to-market firms. He motivates the study by noting that, while a high book-to-market investment strategy, on average, yields positive abnormal returns, the positive returns are generated by only 44% of the sample firms. As such, value investors must tolerate normal or poor returns from a large fraction of the firms in their portfolio. Although a skewed return distribution is not unique to high book-to-market stocks, a trading

strategy that identifies eventual strong value stocks provides more descriptive evidence on the book-to-market puzzle.

To shift the distribution of returns earned by a value investor, the paper describes a trading heuristic based on nine fundamental signals about the financial condition of value firms. The nine signals are based on easily computed measures of profitability, financial leverage / liquidity, and operating efficiency, and each firm's signal realization is classified as either "bad" or "good". Value firms with eight or nine good signals are classified as strong value firms (high F-SCORE), and firms with zero or one good signals are classified as weak value firms (low F-SCORE).

Piotroski suggests that strong value firms will experience improved future firm performance and stock returns relative to weak value firms. The data presented are consistent with this prediction. The mean annual market-adjusted return to a strong value investor is 7.5 percent greater than the mean return to a broad-based value investor. Further, because weak value firms, on average, generate negative abnormal returns, a zero investment portfolio that buys strong value firms and shorts weak value firms generates a 23 percent annual return. These findings are robust to controlling for previously documented trading strategies as well as some common proxies for systematic risk. Within the portfolio of high book-to-market firms, the author presents data consistent with the mean returns to fundamental analysis being concentrated in small and medium sized firms, companies with low trading volume and firms with no analyst following.

This discussion focuses on two main questions. First, does the author's analysis of the data provide convincing evidence that a simple accounting-based trading strategy, when applied to high book-to-market firms, generates substantial abnormal returns? Second, in the absence of

a plausible and rejectable alternative hypothesis to market efficiency, can empirical work on pricing anomalies advance our understanding of pricing behavior?

## *2. Interpreting the Returns to a Strong Value Investing Strategy*

Perhaps more than most research areas, the interpretation of the anomaly and trading strategy literature is heavily influenced by the strong and diffuse priors held by the academic community. The long-standing notion of market efficiency has been increasingly challenged in recent years. Supporters of efficiency react to trading strategies that generate substantial abnormal returns year in and year out over long time periods by attempting to uncover the missing piece of the puzzle (such as risk-adjustment, sample selection biases, or implementability) that will return economic order to the perceived behavior of market participants. Others more willingly accept that these studies have been carried out very carefully, and that economics-based explanations for these empirical regularities have been exhausted. Thus, they conclude that, while markets tend to impound information rationally much of the time, investors occasionally, or even often, systematically ignore information when setting prices. Because of the strength of the priors on both sides, updating occurs slowly.

The comments by conference participants reflected this diversity of priors. The authors' task is two-fold if individuals with strong priors about market efficiency are to be persuaded that a strong value trading strategy successfully separates winners from losers. First, the paper must convincingly document that the returns to a strong value trading strategy are abnormal and achievable. On this note, several conference participants felt that important empirical issues must be resolved before concluding that the results cannot be explained by rational economic behavior. Second, the paper must provide a plausible and rejectable alternative hypothesis to

market efficiency that is consistent with the observed pricing behavior. I address the former task in this section and the latter task in Section 3.

Interestingly, a contributing factor to the skepticism of some conference participants is the size of the abnormal returns generated by the strong value trading strategy, as well as the robustness of these returns to using a wide variety of alternative performance signals. In particular, the magnitude of the abnormal returns generated by this trading heuristic is striking when compared to the realized returns generated by more sophisticated investing strategies employed by hedge funds and mutual funds. Examples of investing strategies employed by hedge funds include convertible arbitrage, distressed securities, market timing, merger arbitrage, as well as value and growth investing (Liang [1999]).

Many studies document that, on average, mutual fund managers do not outperform the market (e.g., Malkiel [1995]; Carhart [1997]). Recent research finds that hedge fund managers generally do, on average, generate positive abnormal returns, although the magnitude of these returns is considerably smaller than the abnormal returns documented in this paper for a strong value trading strategy. For example, Brown, Goetzmann, and Ibbotson [1999] find that while the raw returns for off-shore hedge funds between 1989 and 1995 lagged the S&P 500 by about 3% annually, the risk-adjusted annual return for these funds was a positive 5.7%. Similarly, Liang [1999] documents that average raw returns for a broad sample of offshore and onshore hedge funds from 1994 to 1996 exceeds the return on the S&P 500 by 3% annually, and the median annualized risk-adjusted return for 16 different funds styles is 6.4%. By comparison, the average annual abnormal return to a strong value investing strategy is 13.4%, while the return to a zero investment (high - low) strategy is 23% annually.

Presumably, a primary objective of fund managers is to invest their capital in assets that yield the highest risk-adjusted returns. In this light, it is puzzling that the returns earned by professional hedge fund and mutual fund managers are not at least as great as the returns documented by the academic community for a simple trading heuristic. Also, given the substantial margin by which a strong value trading strategy appears to beat the strategies currently being employed by hedge funds and mutual funds, it is surprising that the abnormal returns to the strategy have persisted for many years without being competed away. The author argues that an advantage of the strong value trading strategy is that it does not require complex methodologies or large amounts of historical financial information. Ironically, this "advantage" is one of the most troubling aspects of the analysis.

To explore the dynamics of a strong value investing strategy, conference participants questioned why high book-to-market firms are the appropriate sample on which to test this trading heuristic. Specifically, what theory predicts a relation between this simple trading strategy and high book-to-market firms, and why wouldn't this heuristic be expected to work in the general population of firms. The author presents no theory, but asserts that, "In contrast [to growth stocks], the valuation of value stocks should focus on recent changes in firm fundamentals...and an assessment of these characteristics is most readily accomplished through a careful study of historical financial statements." However, this statement provides no insight into why this particular trading strategy based on easily computed performance metrics is expected to generate substantial future abnormal returns. A vast body of literature documents that investors can process very complicated financial statement information (e.g., pensions, taxes, signaling with respect to capital structure decisions, cash payout decisions, earnings management and footnote disclosures). Further, the fact that some extant studies provide

descriptive evidence on the market's failure to process particular financial signals is not an appropriate justification for ex ante predictions about ad hoc trading strategies.

The vague justification offered for this trading heuristic and the sample on which it is tested begs the question: What are the frictions in this sample that prevent investors from correctly impounding even the most transparent accounting signals? The paper offers two non-mutually exclusive possibilities. The first is that the firms in this sample are "neglected" by analysts and are "plagued by low levels of investor interest". However, it would be useful to better understand what is meant by "low level of investor interest" and how it is expected to influence the process by which information is impounded in prices. Why wouldn't investors be interested in high book-to-market firms, given that they are apparently undervalued to such a large extent? Presumably investors who hold these high book-to-market stocks are very interested in their performance, so much so that one has to wonder why or whether they would be willing to sell them at such low prices. Is it possible that the low trading volume and illiquidity in these stocks would make it difficult to cost effectively implement the proposed trading strategy? Further, while lower analyst following or investor interest may indicate less or noisier information about high book-to-market firms, it is not clear how or why this lower level of information leads to *biased* stock prices. The second explanation offered to motivate the proposed trading strategy relates to pricing bias and is based on behavioral arguments about market overreaction and underreaction. I address these arguments in greater detail in Section 3.

As an alternative explanation for the paper's findings, several conference participants expressed concerns about appropriate risk adjustment in the computation of abnormal returns. This issue is particularly important given the well-documented problems with parametric test statistics in long-window return studies (Kothari and Warner [1997]; Barber and Lyon [1997]).

The author noted that, in general, a simple risk adjustment explanation based on expected returns is unlikely to fully explain the results. Specifically, intuition suggests that the expected returns for value firms with strong signals about future financial performance should be lower, not higher, than the expected returns for weak value firms. However, to make the findings more convincing, it was suggested that risk-adjusted long window returns based on a more sophisticated technique, such as Fama and French [1993] or Barber and Lyon [1997], could provide a better control for expected returns than the market-adjusted returns used in the study. A related point is the use of the value-weighted market index to compute adjusted returns. Given that high book-to-market firms tend to be relatively small, an equal-weighted market return that gives more weight to smaller firms would seem more appropriate.

In addition to discussing appropriate models of expected returns, conference participants also offered several suggestions to alleviate concerns about an unidentified omitted risk factor. One possibility is to document that the trading strategy's success in high book-to-market firms is unique, in that the trading heuristic does not generate abnormal returns in other samples, such as a random sample of firms, or low book-to-market firms. Another possibility is to document that the returns to this strategy do not persist indefinitely or that positive abnormal returns cannot be generated if the strategy is carried out using performance signals that are sufficiently stale. A further check on the results could stem from recognizing that some firms enter the value stock sample multiple times with differing prospects for future performance (i.e., different F-SCORES). Using these firms as their own control sample, it would be interesting to document that abnormal returns are dependent upon F-SCORE values and are not firm specific. In any event, risk-adjustment remains an important and unresolved issue in this study given its emphasis



on long window returns and the known problems with estimating expected returns for small firms (see Fama [1998]).

Conference participants requested additional information about the influence of data restrictions and CRSP and Compustat selection issues on the findings. These issues are important for at least two reasons. First, it is important that sufficient detail is provided so that future researchers can replicate these results. Second, because of the skewness in returns, particularly in small or poor-performing firms, minor test misspecification or data truncation can lead to serious inference problems in tests of rational pricing (Kothari, Sabino, and Zach [2000]).

A data truncation issue that arises in this study is: How many sample firms delist during the return measurement period, and is it appropriate to set the delisting returns equal to zero? Although this point is not addressed in the paper, the author noted at the conference that incorporating a negative delisting return for firms that delist due to poor performance would likely strengthen the results because the returns for weak value firms would be even more negative than in the reported results. However, this explanation is not completely satisfactory for two reasons. First, firms delist for many reasons including mergers, takeovers and bankruptcies, and these events have very different implications for delisting returns. Second, as illustrated in Section 3 below, the negative abnormal returns to weak value firms cause the largest problems with respect to the behavioral intuition proposed by the author to explain his results. Other data questions posed by conference participants include: How are outliers treated (e.g., firms with negative book value of equity), and why are there many more strong value firms than weak value firms, and is this related to CRSP and Compustat selection issues?

### *3. Interpreting Trading Strategies without Rejectable Alternative Hypotheses*

An important contributing factor that has slowed the acceptance of the anomaly and trading strategy literature is the limited theory that supports the empirical work. In this regard, several conference participants questioned whether the findings in this paper are consistent with a well-defined behavioral pricing theory. In his concluding remarks, the author challenges future research to determine whether the "market behavior in this paper equates to inefficiency or is the result of a rational pricing strategy". However, the returns documented for the proposed trading strategy are anomalous only if they reflect a market inefficiency. Therefore, the distinction between market efficiency and market inefficiency is not only within the scope of the paper, but is essential to assessing and interpreting the paper's findings. A complete explanation for the returns earned by the proposed trading strategy must articulate why investors first overreact to poor performance for all value firms, but then underreact to good news signals for strong value firms, while at the same time, underreact to bad news signals for weak value firms.

Although the paper makes no *ex ante* predictions based on theory, the concluding section suggests that the observed pricing behavior is consistent with the "life cycle hypothesis" advanced by Lee and Swaminathan (2000) where continued poor performance leads to extreme pessimism that may take time to overcome. According to this conjecture, the market fails to respond to good news signals in a timely fashion and strong value firms generate positive abnormal returns as the market slowly incorporates this information into prices. However, this hypothesis is deficient in its ability to explain the pricing behavior documented in this paper. Specifically, it does not offer an explanation for the negative abnormal returns earned by weak value firms. That is, why does the market now appear to *underreact* to the bad news signals about weak value firms given that it had previously *overreacted* to poor performance for all value firms?

The lack of ex ante theory is even more troubling given that one might reasonably argue that the predictions could be opposite those made in the paper. In particular, it is not clear that high book-to-market firms with "good" financial signals should generate positive abnormal returns, and that high book-to-market firms with "bad" signals should generate negative abnormal returns. To motivate a contrasting prediction, note that several prior studies document that high book-to-market firms outperform low book-to-market firms in terms of future returns. Assuming this return differential cannot be explained by differences in risk, a common competing explanation is that investors become overly pessimistic (optimistic) about high (low) book-to-market firms, and drive these stock prices below (above) their intrinsic value. If overreaction to good and bad news is the behavioral paradigm that explains book-to-market anomalies, one might expect that high book-to-market investors would overreact positively to strong signals and overreact negatively to weak signals. This prediction suggests that strong value firms would generate lower future returns than weak value firms, contrary to the predictions and findings in this paper.

To better understand the behavioral aspects of high book-to-market investors, a more thorough analysis of the sample firms would have been helpful. In particular, conference participants suggested partitioning firms based on how they came to be in the high book-to-market sample and how long they had been in the sample, as a way to get at expectations about under- and over-reaction. For example, the high book-to-market sample presumably contains weak firms getting weaker, weak firms that have been weak for some time, and weak firms that are improving. What triggers investors to stop overreacting to bad news and start underreacting to bad news? At what point do investors start underreacting to good news? It would be helpful to

understand the historical pricing behavior of these firms as a function of the performance signals examined in this study.

In light of these issues, it is constructive to consider the following general question: Is it informative to identify an anomalous trading strategy without identifying a plausible and rejectable alternative hypothesis to market efficiency? Addressing this point, Ball [1992] emphasizes that, "It is important to view market inefficiency as a theory to be tested against the evidence, not as the residual claimant on the evidence." Fama [1998] notes that many of the hypotheses put forward as an alternative to market efficiency perform well on the anomalies they are designed to explain, but that their performance with respect to other anomalies is "embarrassing". He goes on to argue, "That dredging for anomalies is a rewarding occupation is suggested by the fact that the anomalies literature shows so little sensitivity to the alternative hypothesis problem. The same authors, viewing different events, are often content with overreaction or underreaction, and are willing to infer that both warrant rejecting market efficiency."

This fundamental problem with anomalies research that is not firmly grounded in theory is evidenced by the asymmetry in the documented returns to trading strategies. Studies of trading strategies that are not based on theory are published only if abnormal returns can be documented. Rather than providing evidence in support of a pricing theory, these studies are marketed as descriptive evidence of a "puzzle", to be resolved by future research. By contrast, when trading strategies are tested within the context of a well-defined alternative hypothesis to market efficiency, findings are potentially informative regardless of whether they support or refute the theory. This latter approach will result in a less biased assessment of the ability of the efficient market hypothesis to explain returns. One way to increase the believability and generalizability

of the anomalies literature is to develop fundamental theories and investigate frictions that drive these trading strategies so that researchers can make ex ante predictions about where to look and test for this behavior or friction. However, in the absence of rejectable alternative hypotheses that reduce the asymmetry with which trading strategies are documented, there is the risk of making a serious Type I inference error - rejecting the null hypothesis of market efficiency when it is true.

#### *4. Conclusion*

Piotroski [2000] finds that a simple accounting-based fundamental analysis strategy, when applied to a broad sample of high book-to-market firms, shifts the distribution of historical returns earned by an investor. The magnitude of abnormal returns earned by this strategy is large, and skeptics at the conference proposed explanations for these findings that include sample selection biases, implementability, or appropriate adjustments for risk. While the above points are common criticisms of the anomalies literature, a thorough treatment of these issues is critical to assessing whether the paper successfully documents a pricing "puzzle". Further, as pointed out in Ball [1992], "The conclusion that markets are inefficient emerges from failing to reject a specific inefficiency hypothesis, not by a process of eliminating all other known explanations for the evidence." The author does not offer a rejectable ex ante theory to motivate a strong value trading strategy, but suggests that the findings are consistent with a loose behavioral theory of under- and over-reaction. However, the fact that the paper's conjectures are able to explain only some of the results, while leaving other results unexplained, indicates that future research is necessary to understand the pricing behavior that underlies high book-to-market firms' stock returns.

## REFERENCES

- BALL, R. "The Earnings-price Anomaly." *Journal of Accounting and Economics* (June/September 1992): 319-345.
- BARBER, B. AND J. LYON. "Detecting Long-run Abnormal Stock Returns: The Empirical Power and Specification of Test Statistics." *Journal of Financial Economics* (March 1997): 341-372.
- BROWN, S.; W. GOETZMANN; AND R. IBBOTSON. "Offshore Hedge Funds: Survival and Performance 1989-95." *Journal of Business* (January 1999): 91-117.
- CARHART, M. "On Persistence in Mutual Fund Performance." *Journal of Finance*. (March 1997): 57-82.
- FAMA, E. "Market Efficiency, Long-term Returns, and Behavioral Finance." *Journal of Financial Economics* (September 1998): 283-306.
- FAMA, E. AND K. FRENCH. "The Cross-section of Expected Stock Returns." *Journal of Finance* (June 1992): 427-465.
- FAMA, E. AND K. FRENCH. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics* (February 1993): 3-56.
- KOTHARI, S. P. AND J. WARNER. "Measuring Long-horizon Security Price Performance." *Journal of Financial Economics* (March 1997): 301-339.
- KOTHARI, S.P.; J. SABINO; AND T. ZACH. "Implications of Data Restrictions on Performance Measurement and Tests of Rational Pricing." Working Paper, Massachusetts Institute of Technology, 2000.

LAKONISHOK, J.; A. SHLEIFER; AND R. VISHNY. "Contrarian Investment, Extrapolation and Risk." *Journal of Finance* (December 1994): 1541-1578.

LEE, C. AND B SWAMINATHAN. "Price Momentum and Trading Volume." *Journal of Finance* (October 2000).

LIANG, B. "On the Performance of Hedge Funds." *Financial Analysts Journal*, (July/August 1999): 72-85.

MALKIEL, B. "Returns from Investing in Equity Mutual Funds 1971 to 1991." *Journal of Finance* (June 1995): 549-572.