Review of Accounting Studies, 6, 165–189, 2001

Contextual Fundamental Analysis Through the Prediction of Extreme Returns

MESSOD D. BENEISH* dbeneish@indiana.edu
Indiana University

CHARLES M. C. LEE CL86@cornell.edu
Cornell University

ROBIN L. TARPLEY rtarpley@gwu.edu
Georgetown University

Abstract. This study examines the usefulness of contextual fundamental analysis for the prediction of extreme stock returns. Specifically, we use a two-stage approach to predict firms that are about to experience an extreme (up or down) price movement in the next quarter. In the first stage, we define the context for analysis by identifying extreme performers; in the second stage we develop a context-specific forecasting model to separate winners from losers. We show that extreme performers share many common market-related attributes, and that the incremental forecasting power of accounting variables with respect to future returns increases after controlling for these attributes. Collectively, these results illustrate the usefulness of conducting fundamental analysis in context.

Keywords: returns prediction, market efficiency, financial statement analysis, volatility, torpedoes, rockets value

1. Introduction

Most fundamental analysis studies involve large sample estimations that span the entire population of firms with available data. However, fundamental analysis as practiced by professional analysts is generally done in a more limited context, typically involving the comparison of a subset of firms with common characteristics. For example, most sell-side financial analysts tend to focus on firms within the same industry, or the same economic sector. Similarly, many buy-side analysts and fund managers specialize in either the universe of "value" stocks, or "growth" stocks.

The concept of contextual fundamental analysis has intuitive appeal. Firms in the same industry have many common operating and risk characteristics that allow for more direct comparison of their financial numbers. The operating and financial constraints of "value" firms, and/or firms in financial distress, can also be expected to differ from those of "glamour" firms, and/or firms with high growth opportunities. For example, in analyzing value stocks we might be more concerned about changes in a firm's capital structure, interest coverage ratio, and short-term liquidity measures. Conversely, in analyzing growth stocks,

*Address correspondence to: Messod D. Beneish, Kelley School of Business, 1309 E. 10th Street, Bloomington, IN 47405-70.

the most important indicators might be those that bear on the sustainability of the firm's current growth trajectory.

In this study, we apply the concept of contextual fundamental analysis to the prediction of extreme stock returns. Every quarter, a subset of the stocks in the universe of U.S. traded securities experience sharp increases or declines in their market price. For example, between 1976 and 1997, size-adjusted returns for stocks in the bottom two percent of the NYSE/ AMEX/ NASDAQ-NMS universe averaged −28% per quarter. Conversely, stocks in the top two percent of the same universe averaged +69% per quarter. Even though gains and losses from holding them can be dramatic, surprisingly little is known about the most salient characteristics of these stocks.

We believe that stocks with a tendency to yield volatile returns offer an interesting economic context in which to apply fundamental analysis. These stocks are of particular interest to professional fund managers and investors.[1] More importantly, they are also likely to share many market and trading characteristics that distinguish them from other securities. Although these characteristics do not necessarily derive from financial statements, they can affect the usefulness of the accounting information we extract from the statements. We examine the proposition that fundamental analysis based on historical accounting variables is more productive, in terms of correlation with future returns, if we first isolate the firms that are likely candidates for sharp price movements (both upward and downward).

Our research design involves two parts. In the first part of our study, we explore the potential usefulness of contextual analysis for extreme performers. In the second part of the study, we develop a procedure for conducting this analysis. We begin by examining the extent to which existing market anomalies, such as the price-to-book effect, the price momentum effect, the post-earnings-announcement drift, and the total accruals effect, are attributable to the extreme performers in the top and bottom 2% of all stocks.[2] We are particularly interested in the differential role that accounting-based variables (such as R&D, profit margins, and capital expenditures) play in returns prediction across extreme and non-extreme firms.

Our results show that variables with predictive power for returns differ sharply across extreme and non-extreme firms. Only three variables are consistently useful predictors of future returns for both extreme and non-extreme firms: total accruals (ACCRUAL), earnings surprises (CHGEPS), and a capital expenditures measure we develop (CAPX). Other variables are useful in only the extreme sample or the non-extreme sample. For example, changes in gross margin (GMG), the price momentum variable (FRTN6), and the firm size variable (SIZE) are correlated with future returns only in the extreme sample. An indicator variable for losses over the past 4 quarters (LSY), and the book-to-market ratio (B/P) forecast returns only in the non-extreme sample.

Still other variables actually operate in opposite directions for extreme and non-extreme firms. For example, R&D expense and sales growth are both favorable returns indicators (positively correlated with returns) among extreme performers, but they are negatively correlated with returns for non-extreme firms. Collectively, these results show that, *if* we can predict extreme performance with perfect foresight, we should use a different set of forecasting variables to separate expected winners from expected losers.

In the second part of the paper, we examine the efficacy of a two-stage procedure for predicting extreme return performance. In the first stage, we estimate a model that identifies extreme performers relative to a control group. This step provides a context for our fundamental analysis. In the second stage, within the subset of predicted extreme performers, we contrast extreme winners ("rockets") with extreme losers ("torpedoes"). We find that, relative to a one-stage direct forecast of future returns, this two-stage procedure appears to be incrementally useful, generating significantly higher pseudo r-squares at both stages.

More importantly, we show that while market-based variables are useful for identifying extreme performers, accounting-based fundamental variables are more useful in separating losers from winners among the subset of predicted extreme performers. Overall, we find that extreme performers are younger, lower priced, and smaller firms with more volatile returns, higher trading volume, and more positive price momentum. Among predicted extreme performers, losers are larger, have lower sales growth, lower price momentum, greater declines in gross margin, more negative recent earnings surprises, more positive total accruals, and higher capital expenditures. Taken together, extreme losers appear to be over-extended growth firms, with slowing sales growth, shrinking margins, but significant levels of new capital investments.

In out-of-sample tests, we find that zero net-investment trading strategies based on our two-stage prediction technique yield substantial size-adjusted returns over the next twelve months. Predicted extreme winners earn substantially higher returns than predicted extreme losers. This effect persists even after controlling for B/P, Size, Price Momentum, Earnings Surprise, and Total Accruals. Approximately one third of the abnormal returns occur around the next four earnings release dates, suggesting at least a portion of these returns are likely due to market misperceptions of future earnings prospects rather than risk-based explanations.

The remainder of the paper is organized as follows. In the next section, we discuss this study in the context of related literature in accounting and finance, and provide motivation for our explanatory variables. In Section 3, we describe our research methodology and sample firms. In Section 4 we present our empirical results, and in Section 5 we summarize our findings and discuss their implications.

2. Related Literature and Explanatory Variables

2.1. Relation to Prior Research

This study is related to three streams of research in accounting and finance. First, our approach is similar to studies that have examined the distinguishing characteristics of firms experiencing unusual events. Second, it is an extension of the studies in financial economics that suggest future returns are predictable. Third, it is related to the growing literature in accounting on the usefulness of historical financial information in predicting significant future events. In this section, we discuss how this study is related to recent work in both finance and accounting. In the course of this discussion, we also introduce and motivate our explanatory variables.

Our approach is related to work that have studied and sought to predict unusual events. For example, researchers have estimated models to predict mergers (Palepu, 1986), bankruptcy (Altman et al., 1977; Zmijewski, 1994), audit qualifications (Dopuch et al., 1987), and earnings manipulation (Beneish, 1997, 1999a). A subset of these studies has used the estimated probabilities from these models to investigate whether strategies using the classification results of their models yielded economically significant returns.

In particular, our paper is similar in spirit to Reinganum (1988). Using a small sample of 222 firms, Reinganum shows that a combination of nine technical and fundamental variables is useful in identifying extreme winners. In particular, he finds that market winners tend to have low B_P ratios, positive recent earnings news, and increasing price momentum. We build on this result by showing that these extreme winners share many common traits with extreme losers. We find that it is incrementally useful to use a two-stage approach to separate predicted extreme winners from predicted extreme losers. Furthermore, we evaluate the incremental contribution of fundamental analysis in both tasks.

Our study is also related to a number of studies that have examined the predictability of cross-sectional returns. For example, Fama and French (1992), Lakonishok, Shleifer, and Vishny (1994), and Davis (1994) document a relation between book-to-market equity ratio (B/P) and subsequent stock returns. Other studies have also identified variables that are related to subsequent returns, including: firm size (Fama and French, 1992), earnings yield (Basu, 1977), cash flow yield (Chan, Hamao, and Lakonishok, 1991), forecasted long-term earnings growth (LaPorta, 1996), recent earnings surprise (Bernard and Thomas, 1989, 1990), total accruals (Sloan, 1996), and past sales growth (Lakonishok, Shleifer and Vishny, 1994). Our use of accounting information to predict future returns is also in the spirit of Ou and Penman (1989), Holthausen and Larcker (1992), Abarbanell and Bushee (1998), and Beneish (1997).

Our study builds on these prior papers in several respects. First, none of the prior studies focus on the relation between predictive signals and extreme performers. In later tests, we contrast the efficacy of the two-stage approach to returns prediction with the one-stage design in these prior studies. Second, we formally integrate explanatory variables from both accounting and finance, while the prior studies focus primarily on either market-related variables, or signals generated by historical financial statements. Third, we examine the incremental usefulness of accounting variables in a contextual analysis.

On this last point, our research is similar in spirit to Piotroski (2000). In that study, the author focuses on the usefulness of fundamental analysis for firms in the bottom B/P quintile (so called "value" firms). He finds that accounting variables derived from historical financial statements play an important role in separating winners and losers among value stocks. By focusing on the predicted returns for extreme performers, the second-stage of our analysis applies fundamental analysis to high volatility stocks. As we show later, these stocks exhibit primarily growth and glamour characteristics. While both studies point to the usefulness of conducting contextual fundamental analysis, the contexts differ, and as we discuss later, so do a number of key results.

2.2. Variable Definition and Motivation

Our first research objective is to provide a descriptive profile of the firms that subsequently experience a sharp price movement. Given this objective, we begin with a broad set of potential explanatory variables gleaned from prior research. In total, we consider 12 market-based variables and 8 fundamental signals. These 20 variables are listed in Table 1 and are described in more detail on the next page.

Table 1. Variable definitions.

Panel A: Dependent Variable	
RTNQ1	3-month size-adjusted buy-hold returns over the first calendar quarter that begins at least 3 months after the current fiscal quarter end (time t).
Panel B: Firm Characteristics	
SIZE	Decile Ranking of Market Cap as of Dec 31 of the year prior to portfolio formation date
PRICE	Natural log of price just before portfolio formation date
AGE	Age of firm computed as number of months from 'BEGDAT' on CRSP to fiscal quarter end
NUMEST	Number of analysts supplying FY1 forecasts (= 0 if firm not in IBES database)
Panel C: Trading Characteristics	
FRTN6	Prior 6-month size-adjusted buy-hold returns
AVGVOL	Prior 6-month Average Daily Turnover
NASVOL	Equals AVGVOL if firm is traded on Nasdaq, zero otherwise
STDRET	Standard deviation of daily returns over the 250 trading days prior to portfolio formation
MINMAX	The ratio of the highest price to the lowest price over the past 30 trading days
Panel D: Market Multiples	
B/P	The book-to-price ratio computed as total SE/MVE as of the portfolio formation date
S/P	The sales-to-price ratio computed as Net sales/MVE
D/P	The debt-to-price ratio computed as (current liabilities + LT debt)/MVE
Panel E: Fundamental Variables	
SGI	(Sales, t/Sales, $t-1$)
GMG	%Δ in sales − %Δ in gross margin
R&D	R&D expense deflated by total assets, deemed zero if missing
LSY	Indicator (= 1) if loss (EBXI) occurs in most recent 4 quarters
CHGEPS	Earnings surprise ((EPS, q-EPS, q-4)/Price, q-4)
ACCRUAL	Total accruals/Average total asset
CAPX	Capital Exp./Average total asset, where Capital Exp. is Item D30 from Annual Compustat
SLDY	Indicator (= 1) if sales decline in most recent 4 quarters
LSY	Indicator (= 1) if loss (EBXI) occurs in most recent 4 quarters

Note: CAPX was not available quarterly, and was obtained from Annual Compustat. All other accounting numbers are based on rolling quarterly statements. Year t refers to data from the most recent 4 quarters. All %Δ's calculated as follows: $[X_t - ((X_{t-1} + X_{t-2})/2)]/((X_{t-1} + X_{t-2})/2)$.

Panel A of Table 1 presents the dependent variable: size-adjusted returns in a subsequent quarter (RTNQ1). This variable is the buy-and-hold return for a given firm in a future quarter, minus the average return for firms in the same size decile over that quarter. To ensure the accounting information we use is available by the portfolio formation date, we compute RTNQ1 over the first calendar quarter that begins at least 3 months after each fiscal quarter end.

Our independent variables derive from two main streams of research. First, from the literature on market volatility, we glean a number of variables hypothesized to predict extreme performance (i.e., higher moments of expected returns). Second, from prior literature on market pricing anomalies, we derive a set of variables hypothesized to predict the mean of expected returns. In the following discussion, we group these variables as either market-based signals, or fundamental signals. These two groupings are intended to highlight the incremental contribution of accounting variables to the two tasks at hand.

2.2.1. Market-based Signals

The purpose of these market-based signals is to control for general firm characteristics and recent trading patterns that could be early indicators of large imminent price moves. Panel B of Table 1 presents four general firm characteristics. SIZE is the decile ranking of a firm's market capitalization based on NYSE cutoffs, as of December 31 of the prior year. PRICE is the natural log of price just before portfolio formation. NUMEST is the number of analysts supplying a one-year-ahead earnings forecast (assumed to be zero if a firm is not in the IBES database). AGE is defined as the number of months from the beginning listing date on CRSP to the fiscal quarter end. The first three variables are highly correlated, and proxy for liquidity and information risk (Barry and Brown, 1984; Fama and French, 1992). We expect smaller firms to be more volatile and to earn higher expected returns. The AGE variable is motivated by prior research suggesting higher incidences of financial statement fraud and financial distress among young growth firms. As these financial conditions come to light in the future, these firms' experience large price movements (Beneish, 1999b; Seyhun and Bradley, 1997).[3]

Panel C presents five variables associated with recent trading activities in each stock. FRTN6 is the size-adjusted buy-and-hold return in the six months prior to portfolio formation, AVGVOL is the average daily turnover ratio over the prior six months, and NASVOL is AVGVOL*I, where $I = 1$ for Nasdaq stocks and 0 otherwise. These three variables are motivated by the pricing anomaly literature. Jegadeesh and Titman (1993) show that firms with higher (lower) FRTN6 earn higher (lower) returns over the next 12 months. Lee and Swaminathan (2000) show firms with higher (lower) trading volume (AVGVOL) exhibit glamour (value) characteristics, and earn lower (higher) returns in subsequent months. In measuring turnover, Nasdaq stocks are treated separately because trading volume from dealer-driven markets including some double counting of public orders (Lee and Swaminathan, 2000).

We also include three return volatility measures nominated by BARRA's E3 U. S. Equity model (BARRA, 1998). STDRET is the standard deviation of daily returns in the 250 trading days prior to portfolio formation. MINMAX is ratio of the highest daily closing price to the lowest closing price over the past 30 trading days. PRICE is the natural log of price just before portfolio formation. We expect a higher frequency of extreme performers among low PRICE, high STDRET and high MINMAX firms.

Finally, Panel D describes three market-based multiples that we include in our model. D/P, the debt-to-market ratio, is nominated by prior studies on volatility (Beaver, Kettler and Scholes, 1970; Christie, 1982). We expect more levered firms to be more volatile and possibly earn higher expected returns. B/P and S/P are the book-to-price and sales-to-price ratios respectively. Both variables are nominated by the pricing anomalies literature as variables that are positively correlated with future returns (e.g., Fama and French, 1992; O'Shaughnessy, 1998). Interpretations differ among academics as to the explanation for these higher returns. For purposes of this study, these variables control for differences in expected returns not captured by the fundamental signals.

2.2.2. *Fundamental Signals*

Panel D presents 8 fundamental signals. This is clearly not an exhaustive list of potentially useful variables, but these are representative of variables that span much of the prior literature on financial analysis. The first variable, the rate of sales growth over the past year (SGI), is nominated by Beneish (1999a), which shows that earnings manipulators tend to have high sales growth. The second variable (GMG) is nominated by the results reported by Abarbanell and Bushee (1998) and Lev and Thiagarjan (1993). According to these studies, an increase in either measure is an indication of deteriorating earnings quality.

Two other variables associated with higher expected returns are CHGEPS and ACCRUAL. CHGEPS is measure of the earnings surprise from the most recent quarter. The post-earnings announcement drift literature (e.g., Bernard and Thomas, 1989, 1990) reports that high CHGEPS firms earn higher subsequent returns. ACCRUAL is total accruals scaled by average total assets. Sloan (1996) show that the accrual-related component of earnings is less persistent, and that firms with more positive (income increasing) accruals earn higher subsequent returns.

We also include a measure of research and development expenditure intensity (R&D), calculated as total R&D expenditures divided by total assets. Chan, Lakonishok, and Sougiannis (1999) report that R&D intensity is associated with return volatility. They also show that, within the set of growth stocks, R&D intensive firms tend to outperform stocks with little or no R&D. We examine the incremental usefulness of this metric for predicting extreme return performance, and for separating winners from losers.

We introduce a new variable, CAPX, measured as total capital expenditures divided by average total assets. Continued capital investments are important to growth firms. These investments reduce current free cash flows, but are indicative of new projects with uncertain payoffs. The issue is whether the market, on average, prices these payoffs correctly. If the market over- (under-) weights the expected payoffs from capital projects, this variable would be negatively (positively) correlated with future returns.

Finally, we use two indicator variables (LSY and SLDY) to capture possible asymmetry in the case of loss firms (e.g., Hayn, 1995), or firms that have experienced sales declines (e.g., Dietrich and Kaplan, 1982). LSY takes the value of 1 if total earnings-before-extraordinary-item is negative in the four quarters just ended, and 0 otherwise. SLDY takes on the value of 1 if sales declined over the past four quarters, and 0 otherwise.

Taken together, these 20 variables include most of the more common signals discussed in the finance and accounting literature, as well as several new variables. Ours is by no means an exhaustive list of potentially useful metrics. However, these variables should serve as a good basis for developing a descriptive profile of extreme performers, and for evaluating the potential of contextual analysis.

3. Research Methodology

Our sample consists of all firms in the CRSP and merged Compustat (PST, Full Coverage, and Research) universe. To ensure sample firms have sufficient market liquidity, we require each to have a stock price of at least $5 on the portfolio formation date. The time period covered by our data is 1/76 to 12/98. Our first portfolio formation date is 7/77 and our last date is 10/97. After eliminating ADRs, REITs, and closed-end funds, we are left with 59,589 firm-quarter observations for 4,114 firms.

From this original sample, we create an estimation sample and a holdout sample using a time-period and firm-size matching procedure. First, all 59,589 observations are sorted by calendar quarter. Next, each firm-quarter observation is sorted on the basis of the firm's market capitalization as of the previous December 31 (the SIZE variable). We then create two size-matched samples by assigning firms within each quarter to the two samples in the following manner: 1, 2, 2, 1, 1, 2, 2, ... This procedure ensures that we have an equal number of observations in each sample for each calendar quarter, and that these samples are closely matched in terms of firm size characteristics.

Next, we create a categorical dependent variable using size-adjusted returns from the "target quarter." The "target quarter" for each observation is defined as the first calendar quarter that begins at least 3 months after the current fiscal quarter end. The 3-month lag ensures accounting information from the current fiscal quarter is available before the return accumulation period. Firms whose size-adjusted returns are in the bottom (top) 2% during the target quarter are deemed to be extreme loser (winner) firms. Firms whose size-adjusted returns rank in the middle 96% are control firms.

4. Empirical Results

4.1. Univariate Statistics

Table 2 reports summary statistics for these variables based on the 29,795 firm-quarters in the estimation sample. The results for the extreme losers (winners) are reported as Losers (Winners). The results for the remaining firms are reported under the heading Control. The T-statistic (Z-statistic) is from a two-tailed test of a difference in mean (median) between groups. We first compare the results between each of the extreme groups to the control group. In the last two columns, we compare the mean (median) of extreme losers to extreme winners. The number of $+(-)$ signs indicate the direction of the relation and statistical significance at the 1% or 5% level. N is the number of observations in each group.

Panel A shows that out of a total of 27,995 observations, 580 (610) observations are classified as extreme Winners (Losers). Extreme winners experienced a mean size-adjusted return for over the target quarter of 68.9%, while the extreme losers experienced a mean size-adjusted return of -27.8%. In the Control sample, the average size-adjusted return is -0.7% over the target quarter.

Panels B and C show that, compared to the control group, extreme performers tend to be younger (AGE), smaller (SIZE) and lower priced (PRICE) firms, with more volatile returns (STDRET and MINMAX), higher recent trading volume (AVGVOL and NASVOL), and lower analyst coverage (NUMEST). These results suggest that such firms are at greater risk of experiencing extreme price moves. Several of these variables also seem to have some ability to distinguish between extreme winners and losers. Specifically, compared to Losers, the Winners tend to be smaller in size, have lower analyst coverage, and higher return volatility.

Panel B also suggests that prior price momentum (FRTN6) helps to discriminate between extreme losers and extreme winners. Consistent with the price momentum effect (Lee and Swaminathan (2000) and Jegadeesh and Titman (1993)), extreme losers (winners) have experienced more negative (positive) returns in the past 6 months. The size-adjusted return over the past six months for Winners (Losers) is also significantly higher (lower) than for the control group. The results suggest that FRTN6 may be useful in distinguishing winners from losers.

Panel D shows that the market multiples are not strong factors in either identifying extreme performers or returns forecasting. We observe some evidence that extreme winners have somewhat higher S/P and D/P ratios than the control group. The winners also have higher S/P ratios than the losers. In general, however, these market multiples are only of marginal importance.[4]

Panel E shows a number of fundamental variables are likely to be useful in our subsequent analysis. Extreme performers tend to be firms with higher sales growth (SGI) and higher research and development expenditures (R&D). They are also much more likely to have experienced losses (LSY) relative to the control firms. In fact, over 20% of the extreme performers reported losses in the preceding four quarters, relative to just 12.3% of the control firms.

The last two columns of Panel E report results that might be useful in separating winners from losers. These results show that compared to Losers, the Winners have less margin erosion (GMG), and generally reported better earnings in the most recent quarter (CHGEPS). Winners also have much more income-decreasing accruals (ACCRUAL) and lower capital expenditures (CAPX). In fact, the extreme losers have even more positive accruals and higher capital expenditures than the control group.

The picture that emerges from Table 2 is that extreme performers are young growth stocks with "glamour" characteristics. They are smaller firms with higher sales growth, lower valuations ratios, higher trading volume, lower analyst coverage, and are more likely to have reported a loss in the trailing 12 months. Taken together, this set of facts characterize firms bearing greater risk of a sharp price move, and help to define the economic context of our analysis.

4.2. Future Returns to Individual Signals for Extreme and Non-Extreme Firms

To further motivate the case for a contextual analysis, we examine the differential role played by these signals in predicting returns for extreme and non-extreme firms. Table 3 reports the results of a pooled regression based on the 29,795 firm-quarters in the estimation sample (1977–1997). The dependent variable is the one-quarter-ahead size-adjusted returns (RTNQ1). The independent variables are the eight fundamental variables described in Tables 1 and 2, plus the book-to-market ratio (B/P), price momentum (FRTN6), and firm size (SIZE).

Models A and B report a direct regression of the explanatory variables on RTNQ1. These two models are similar to the regressions conducted in prior studies. Models C and D includes a set of interaction terms ($I * y$), in which each variable (y) is multiplied by an indicator variable (I) that equals one if the firm is an extreme performer in the target quarter,

Table 2. Summary statistics for winners, losers, and control groups.

This table presents summary statistics for firms in the extreme winner portfolio (Winners), extreme loser portfolio (Losers), and a control group (Control). Extreme winners (losers) are defined as firm ranked in the top (bottom) two percent in terms of size-adjusted returns in the subsequent calendar quarter. The remaining firms are in the control group. The variables are defined in Table 1. The sample period is 1/76 to 12/98. These statistics are based on the 27,995 firm-quarters in the estimation sample (i.e., they exclude observations from the holdout sample). The T-statistic (Z-statistic) are from two-tailed tests of a difference in mean (median) between these groups. N is the number of observations in each group. − or −− (+ or ++) indicates the statistic is negative (positive) and significant at the 5% or 1% level.

	Winners			Control			Losers			Statistical Tests: Winners vs. Control		Statistical Tests: Losers vs. Control		Statistical Tests: Winners vs. Losers	
	N	Mean	Median	*N*	Mean	Median	*N*	Mean	Median	*T*-stat	*Z*-stat	*T*-stat	*Z*-stat	*T*-stat	*Z*-stat
Panel A. Dependent Variable															
RTNQ1	580	0.689	0.625	26,805	−0.007	−0.012	610	−0.278	−0.281	++	++	−−	−−	++	++
Panel B. Firm Characteristics															
SIZE	580	5.474	5.000	26,805	7.078	7.000	610	6.482	7.000	−−	−−	−−	−−	−−	−−
PRICE	580	2.492	2.397	26,805	2.863	2.862	610	2.578	2.482	−−	−−	−−	−−		
AGE	580	148.94	114.00	26,805	214.14	168.00	610	155.91	125.50	−−	−−	−−	−−		
NUMEST	580	3.283	1.000	26,805	6.164	3.000	610	4.805	2.000	−−	−−	−−	−−	−−	−−

	Winners			Control			Losers			Statistical Tests					
										Winners vs. Control		Losers vs. Control		Winners vs. Losers	
	N	Mean	Median	*N*	Mean	Median	*N*	Mean	Median	*T*-stat	*Z*-stat	*T*-stat	*Z*-stat	*T*-stat	*Z*-stat
Panel C. Trading Characteristics															
FRTN6	580	0.202	0.083	26,805	0.036	−0.004	610	0.075	−0.022	++	++			++	++
AVGVOL	580	0.522	0.298	26,805	0.367	0.224	610	0.507	0.315	++	++	++	++		
NASVOL	580	0.369	0.057	26,805	0.203	0.000	610	0.316	0.000	++	++	++	++		++
STDRET	580	0.036	0.034	26,805	0.027	0.024	610	0.034	0.031	++	++	++	++	++	++
MINMAX	580	1.469	1.213	26,805	1.195	1.158	610	1.257	1.212	++	++	++	++		
Panel D. Market Multiples															
B/P	578	0.680	0.557	28,422	0.697	0.581	607	0.689	0.548						
S/P	580	2.900	1.600	26,805	2.430	1.516	610	2.440	1.407	++				++	
D/P	580	0.928	0.468	26,805	0.796	0.469	610	0.818	0.438	+					
Panel E. Fundamental Variables															
SGI	580	1.206	1.140	28,605	1.149	1.104	610	1.196	1.128	++	++	++	++		
GMG	580	0.243	0.198	28,605	0.219	0.158	610	0.327	0.226		++	++	++	−	
R&D	580	0.029	0.000	28,605	0.014	0.000	610	0.024	0.000	++	++	++	++		
CHGEPS															

Table 3. Future returns to individual signals for extreme and non-extreme firms.

This table presents the results of a pooled regression based on 29,795 firm-quarters in the estimation sample (1977–1997). The dependent variable is the one-quarter-ahead size-adjusted returns (RTNQ1). The independent variables are the eight fundamental variables described in Table 1, plus the book-to-market ratio (B/P), a measure of price momentum (FRTN6), and a measure of firm size (SIZE). Models C and D also feature interaction terms ($I * y$), in which each variable (y) is multiplied by an indicator variable (I) that equals one if the firm is an extreme performer in the target quarter, and zero otherwise. Table values represent estimated coefficients. ***, **, * indicates statistical significance at the 1%, 5%, and 10% level respectively, based on two-tailed tests.

Variable	Model A	Model B	Model C	Model D
Intercept	−.011	.007	.003	.009
Primary Explanatory Variables				
SLDY	−.005	**−.006***	**−.007****	**−.011*****
SGI	**.018*****	**.018*****	−.002	**−.013***
GMG	−.003	−.004	−.000	.002
R&D	.019	.029	**−.054****	**−.043***
LSY	**−.026*****	**−.030*****	**−.033*****	**−.034*****
CHGEPS	**.128*****	**.103*****	**.040***	**.043***
ACCRUAL	**−.141*****	**−.140*****	**−.137*****	**−.127*****
CAPX	**−.086*****	**−.077*****	**−.061*****	**−.059*****
B/P		**.006*****		**.006*****
FRTN6		**.020*****		.004
SIZE		**−.002*****		.000
Interaction Variables				
I*SLDY			−.016	**.048*****
I*SGI			**.211*****	**.453*****
I*GMG			**−.059*****	**−.117*****
I*R&D			**.549*****	**.593*****
I*LSY			.010	.008
I*CHGEPS			**.588*****	**.358*****
I*ACCRUAL			**−.230*****	**−.341*****
I*CAPX			**−.599*****	**−.392*****
I*B/P				−.005
I*FRTN6				.100***
I*SIZE				**−.028*****
Adjusted r-square	0.62%	0.80%	5.5%	6.3%

and zero otherwise. In Models C and D, the non-interactive terms represent the predictive power of the variables in the control group. The interactive terms represent the incremental effect of these variables in forecasting returns among the extreme performers.

Table 3 shows that variables with predictive power for returns differ sharply across extreme and non-extreme firms. In Model D, only three variables are consistently useful predictors of future returns for both extreme and non-extreme firms: total accruals (ACCRUAL), earnings surprises (CHGEPS), and the capital expenditures measure (CAPX). Some variables are useful only for control firms: an indicator variable for losses over the past 4 quarters (LSY), and the book-to-market ratio (B/P). Other variables are useful only among extreme performers: changes in gross margin (GMG), the price momentum variable (FRTN6), and the firm size variable (SIZE).

Still other variables actually operate in opposite directions for extreme and non-extreme firms. For example, R&D expense and sales growth are both strongly favorable returns indicators (i.e., positively correlated with returns) among extreme performers, but they are negatively correlated with returns for non-extreme firms. Collectively, these results show that, *if* we can predict extreme performance with perfect foresight, we should use a very different set of forecasting variables to separate expected winners from expected losers.

4.3. Multivariate Estimation Results

In Table 4, we estimate three sets of multivariate Probit regressions. In the first column, the dependent variable is coded 0 for control firms and 1 for extreme performers (Extreme vs. Controls). In the second column, firms with negative size-adjusted returns in the target quarter are coded 1 and firms with non-negative size-adjusted returns are coded 0 (Losers vs. Winners, Full Sample). In the third column (Losers vs. Winners, Extreme Sample), we limit our sample to the set of extreme firms. In this limited sample, extreme losers are coded 1 and extreme winners are coded 0.

The middle column reports the results of a direct forecast of returns for the full sample, and is closest to the research design in most prior studies. Notice that the estimated coefficients show that the prior pricing anomalies generally replicate in this regression. However, as in prior studies, the pseudo-R^2 is low, at 0.38%.

The first column reports the results of a model for explaining extreme performance. This model shows most of the univariate results continue to hold. Extreme performers are: smaller (SIZE), younger (AGE), lower priced (PRICE), and more volatile (STDRET). They also have higher price momentum (FRTN6), higher turnover (AVGVOL) and higher sales-to-price ratios (S/P).

The accounting variables provide further insights. As a group, extreme firms have higher average sales growth (SGI), but are also more likely to have reported a sales decline over the past 12 months (SLDY). Taken together, these two variables reflect the volatile nature of their revenue stream. Table 4 also shows that the extreme performers have greater R&D expenditures (R&D), are more likely to have reported a loss before extraordinary items (LSY), and have had more positive earnings surprises in the last quarter (CHGEPS). The model has a pseudo-R^2 of 5.2%.

The last column in Table 4 reports the regression of Winners versus Losers in the sub-sample of extreme firms. This regression shows that losers are larger in size (SIZE) and had more negative returns over the past 6 months (FRTN6). None of the other market-related variables are useful in separating winners from losers, but several fundamental variables play a significant role. Specifically, extreme losers reported lower sales growth (SGI), experienced greater declines in their gross margins (GMG), had more negative earnings surprises (CHGEPS), reported more income-inflating accruals (ACCRUAL), spent less on research and development (R&D), and more on capital expenditures (CAPX), than the extreme winners. In short, extreme losers exhibit many of the signs analysts ascribe to weaker performing growth firms. Comparing the second and third columns, it is evident that the winner-vs-loser model explains a much larger proportion of the variation in future returns when it is estimated on the extreme sample.

Table 4. Probit estimation results over the full sample period (1977–1997).

This table reports the regression results of three Probit estimations. In the first column, the dependent variable equals 1 if the firm is an extreme performer, zero otherwise.[a] In the second column, the dependent variable is equal 1 if the firm is a loser (if RTNQ1 < 0), zero otherwise. In the third column, the same winner versus loser regression is estimated, but only for the subpopulation of extreme performance firms. Table values represent estimated coefficients, with p-values reported in parentheses. Variables that have statistical significance at 5% or lower are highlighted with bold type.

Variables	Extreme Firms vs. Control Firms	Losers vs. Winners (Full Sample)	Losers vs. Winners (Extreme Sample)
Constant	−1.72	−0.02	−.084
	(.000)	(.975)	(.811)
General Firm Characteristics			
SIZE	**−.018**	**.010**	**0.098**
	(.022)	**(.018)**	**(.000)**
PRICE	**−.175**	**−.038**	−.037
	(.000)	**(.024)**	(.657)
AGE	**−.0004**	−.000	.000
	(.000)	(.648)	(.438)
Recent Trading Characteristics			
FRTN6	**.107**	−.036	−.138
	(.001)	(.107)	(.062)
AVGVOL	**.196**	**.100**	.146
	(.000)	**(.004)**	(.333)
NASVOL	**−.118**	**−.088**	−.189
	(.018)	**(.007)**	(.177)
STDRET	**6.45**	**5.00**	5.60
	(.000)	**(.000)**	(.120)
MINMAX	−.027	−.016	−.032
	(.345)	(.156)	(.249)
Market Multiples			
S/P	.010	**−.009**	−.002
	(.077)	**(.011)**	(.891)
D/P	−.003	.006	−.057
	(.857)	(.568)	(.274)
Fundamental Variables			
SLDY	**.082**	.024	.016
	(.042)	(.271)	(.882)
SGI	**.229**	−.023	**−.527**
	(.001)	(.596)	**(.010)**
GMG	**−.018**	−.016	**.223**
	(.481)	(.311)	**(.032)**
R&D	**.618**	−.217	−1.01
	(.013)	(.202)	(.115)
LSY	**.098**	**.104**	.147
	(.018)	**(.000)**	(.162)
CHGEPS	**.466**	**−.493**	**−1.37**
	(.046)	**(.002)**	**(.024)**
ACCRUAL	.179	**.747**	**1.357**
	(.249)	**(.000)**	**(.001)**
CAPX	.214	**.301**	**1.062**
	(.219)	**(.003)**	**(.046)**

(*Continued*)

Table 4. (*Continued*)

	Extreme Firms vs. Control firms	Losers vs. Winners (Full Sample)	Losers vs. Winners (Extreme Sample)
Pseudo-R^{2b}	**5.23%**	**0.38%**	**4.42%**
χ^2-Statistic	**538.4**	**223.8**	**98.6**
P-Value	**.001**	**.001**	**.001**
Pseudo-R^2 for fundamental variables only[b]	**1.70%**	**0.24%**	**2.11%**

[a]The estimation sample contains up to 29,795 firm-quarter observations over the period 1977–1997. Individual variables are described in Table 1. Extreme performers refer to firms in either the top or bottom two percent of the distribution of quarterly returns and controls refer to the remaining 96%. Losers (Winners) are firms in the lowest (highest) two percent of the distribution of quarterly returns.

[b]The pseudo-R^2 is equal to $(L_{\Sigma}^{2/n} - L_w^{2/n})/(1 - L_w^{2/n})$ where L_Σ is the log likelihood for the unconstrained Probit model, L_w is the log likelihood with only the constant term in the model (constrained) and n the number of observations (See Maddala (1983, p. 40)). The log likelihood ratio test statistic, equal to -2 times the difference in the log likelihood of the unconstrained and constrained models, is asymptotically distributed X^2, with degrees of freedom equal to the difference in the number of parameters of the two models. The pseudo-R^2s reported in last row are for a model that uses only the eight fundamental variables.

We also use the methodology proposed in Beneish and Harvey (1998) to compare the pseudo R^2s for market-based and accounting-based variables. The last row in Table 4 reports the pseudo R^2 for accounting-based variables only. For this purpose, we defined only the eight fundamental variables as accounting-based, all other variables are deemed market-based. Our tests show that market-based variables are significantly more powerful for distinguishing extreme performers from the control group. Accounting-based fundamental signals play a much bigger role in separating extreme winners from extreme losers.

In sum, our analyses highlight the differential roles played by accounting numbers in identifying risky stocks, and in separating extreme winners from extreme losers. Both the extreme-vs-control model and the winner-vs-loser model have statistical power to discern their target variables. The latter model has much higher explanatory power when conducted on the sample of extreme firms than on the general population. These in-sample results are encouraging, suggesting that extreme performance may be predictable in advance. We turn to out-of-sample results in the next section.

4.4. Out-of-Sample Classification Tests

We evaluate the model's ability to distinguish between winners and losers in a holdout sample that contains up to 29,794 firm-quarter observations over the period 1977–1997. The advantage of our matching process is that the estimation and holdout samples are closely matched in terms of time period and firm size. The disadvantage of this approach is that, strictly speaking, we cannot implement a trading strategy based on this model.

The model's cut-off probabilities for classifying firms either as winners or losers are based on minimizing the expected costs of misclassification (ECM), given by:

$$\text{ECM} = P(E)P_I C_I + [1 - P(E)]P_{II}C_{II},$$

where $P(E)$ is the prior probability of encountering extremes (.04 by construction for extremes [both tails], and .02 for either winners or losers), P_I and P_{II} are the conditional probabilities of, respectively, Type I and Type II errors, and C_I and C_{II} are the costs of Type I and Type II errors. In Panel A of Table 5, a Type I error is defined as the failure to classify an observation as an extreme performer, and a type II error is defined as classifying an observation as an extreme performer when it is not. Cut-off probabilities were chosen for each level of relative costs to minimize, in the estimation sample, the expected costs of misclassification (ECM), as defined in the equation above.

The "correct" relative cost of Type I and Type II errors is, of course, a matter of subjective judgement. We consider relative costs ranging from 10 : 1 to 40 : 1 for the tests in Panel A of Table 5, because these assumptions should span the relevant range for most investors. For example, we report in Table 2 that the average loser's stock declines (on a size-adjusted basis) by 28% in a quarter. Assuming that a typical firm's equity appreciates between 1% and 3% per quarter, it takes 10 to 30 non-losers in an investor's portfolio to offset a single loser in that quarter. As such, investors may reasonably view a type I error as 20 to 40 times as costly as a type II error. In Panel B, we assume a 1 : 1 relative cost in terms of the classification of extreme winners to extreme losers. We base this assumption on the fact that the number of winners and losers used in the sample are equal by construction.

Table 5 reports the estimated probabilities and the classification results for both predictive models. In the first two columns of each panel, we report the models' mean estimated probabilities for the treatment and control firms. These results show that the "extreme-vs-control" model has predictive ability in both the estimation and holdout samples. For example, in Panel A, the mean estimated probabilities for actual extreme performers (.051 to .119) are generally three times larger than the corresponding probabilities for the control firms (.016 and .034). The t-tests and median tests easily reject the null hypothesis that estimated probabilities for extremes and controls are drawn from the same distribution.

Panel A shows that, under a relative cost assumption of 10 : 1, Model 1 correctly classifies 15.5% of the extreme performers, and incorrectly classifies 5.6% of the control firms. In the Holdout sample, this model correctly classifies 16.2% of the extreme firms and incorrectly classifies 6.0% of the control firms. As we increase relative costs, we are able to correctly classify a greater proportion of the extreme firms, but the proportion of incorrectly classified control firms also increases. For example, using a 30 : 1 cost assumption, it is possible to accurately classify 85.3% of the extreme performers, while picking up 60.5% of the non-extreme firms. This level of classification accuracy may be of some interest to investors with a strong aversion to extreme return performance.

Perhaps of greater interest is the success of the "loser-vs-winner" model. Panel B shows that this model has significant ability to separate extreme winners from extreme losers. In the estimation (holdout) sample, this model correctly identifies 61.6% (61.6%) of the losers while incorrectly classifying only 30.9% (30.7%) of the winners. Together with the results in the earlier panel, these findings suggest that a two-stage approach could improve our ability identify extreme winners and extreme losers in advance.

Table 6 examines the mean size-adjusted returns for firms identified by the various models as winners and losers in the holdout sample. As we mentioned, the trading strategy implied by this table can not be implemented, because the holdout sample spans the same time period as the estimation sample. However, these results do provide an indication of how

Table 5. Classification accuracy in estimation and holdout samples.

This table reports classification results for four Probit regression models. The left-side results are for the Estimation Sample, and the right-side results are for the Holdout Sample. In Panel A, we report the mean estimated probabilities from the Extremes versus Controls model, and the percentage of correct (incorrect) classification under various relative cost assumptions. In Panel B, for the Loser versus Winners model, we report the same statistics under a 1 : 1 relative cost assumption (i.e., 50% probability of being either a Loser or a Winner.) In Panel A (B), the reported mean estimated probability is the average ex ante probability that the model assigns to the observations in a given group of being an extreme performer (loser). *** indicates that the difference in mean estimated probabilities is different from zero at the 1% level.

Panel A: Extremes versus controls

	Estimation Sample				Holdout Sample			
Relative Cost Assumption	Mean Est. Probability for Extremes	Mean Est. Probability for Controls	Correctly Classified Extremes	Incorrectly Classified Controls	Mean Est. Probability for Extremes	Mean Est. Probability for Controls	Correctly Classified Extremes	Incorrectly Classified Controls
10 to 1	0.119***	0.034	15.46%	5.62%	0.118***	0.035	16.16%	5.95%
20 to 1	0.068***	0.024	65.21%	35.63%	0.068***	0.024	64.56%	36.00%
30 to 1	0.054***	0.017	87.31%	59.78%	0.054***	0.017	85.22%	60.48%
40 to 1	0.051***	0.016	92.61%	68.05%	0.051***	0.015	90.75%	68.45%

Panel B: Losers versus winners

	Estimation Sample				Holdout Sample			
Relative Cost Assumption	Mean Est. Probability for Losers	Mean Est. Probability for Winners	Correctly Classified Losers	Incorrectly Classified Winners	Mean Est. Probability for Losers	Mean Est. Probability for Winners	Correctly Classified Losers	Incorrectly Classified Winners
1 to 1	0.616***	0.408	64.92%	30.91%	0.616***	0.406	62.76%	30.72%

Table 6. Predicted winners, predicted losers, and future returns.

This table reports mean size-adjusted returns for firms in the holdout sample over the 1977–97 period. Period A reports results for firms classified as "Extreme" and "Non-extreme" by the "Extreme versus Control" model under different relative error cost assumptions. Panel B reports results of a two-stage classification, in which firms classified as Extreme in the first stage are re-classified as either predicted winners or predicted losers using the "Losers versus Winners" model. * (**) [***] indicates rejection of the null that mean returns for the classified and non-classified observations are equal at the 10% (5%) [1%] level based on two-tailed tests.

Panel A: One-stage classification using the "Extremes versus Control" model

Relative Cost Assumption	Classified Extreme?	*N*	Quarter One	Quarter Two	Quarter Three	Next 6 Months	Next 12 Months
10 to 1	Yes	1,892	−0.001	−0.001	−0.003	0.001	−0.002
10 to 1	No	27,902	0.003	0.001	0.002	0.002	0.003
10 to 1	*Difference*		*−0.004*	*−0.002*	*−0.005*	*−0.001*	*−0.005*
20 to 1	Yes	11,057	0.002	−0.003	0.001	−0.001	−0.002
20 to 1	No	18,737	0.003	0.002	0.003	0.005	0.006
20 to 1	*Difference*		*−0.001*	*−0.005*	*−0.002*	*−0.006*	*−0.008*
30 to 1	Yes	11,487	0.003	−0.002	0.002	0.001	0.003
30 to 1	No	18,307	0.002	0.004	0.002	0.004	0.003
30 to 1	*Difference*		*0.001*	*−0.006*	*0.000*	*−0.003*	*0.000*
40 to 1	Yes	20,654	0.003	−0.001	0.003	0.001	0.004
40 to 1	No	9,140	0.002	0.003	0.001	0.002	0.000
40 to 1	*Difference*		*0.001*	*−0.004*	*0.002*	*−0.001*	*0.004*

Panel B: Two-stage classification separating predicted extreme performers using the "Loser versus Winner" model

Relative Cost Assumption	Classified Losers?	*N*	Quarter One	Quarter Two	Quarter Three	Next 6 Months	Next 12 Months
10 to 1	Yes	620	−0.050	−0.010	−0.033	−0.064	−0.108
10 to 1	No	1,272	0.023	0.003	0.012	0.033	0.050
10 to 1	*Difference*		*0.073* ***	*0.013*	*0.045* ***	*0.097****	*0.158****
20 to 1	Yes	5,186	−0.016	−0.014	−0.010	−0.035	−0.062
20 to 1	No	5,871	0.019	0.010	0.010	0.029	0.052
20 to 1	*Difference*		*0.035* ***	*0.024* ***	*0.020* ***	*0.067* ***	*0.114* ***
30 to 1	Yes	8,800	−0.010	−0.011	−0.006	−0.023	−0.039
30 to 1	No	7,788	0.020	0.010	0.012	0.032	0.058
30 to 1	*Difference*		*0.030* ***	*0.021* ***	*0.018* ***	*0.055* ***	*0.097* ***
40 to 1	Yes	12,146	−0.008	−0.009	−0.004	−0.020	−0.032
40 to 1	No	8,508	0.019	0.011	0.012	0.032	0.058
40 to 1	*Difference*		*0.027* ***	*0.020* ***	*0.016* ***	*0.052* ***	*0.090* ***

well the model's classifications correspond to subsequent price performance in a separate size-matched sample.

Table 6 Panel A reports mean returns to firms classified and firms not classified as extreme performers by the Extremes vs. Controls model, under four different relative cost assumptions. The results show that, on average, the subsequent returns of the predicted extreme firms are not distinguishable from those of the control firms. Over the next 12 months, firms classified by this model as extreme performers earn differential size-adjusted returns of around −0.5% (under relative costs of 10:1). This amount is not significantly different from zero.

Panel B reports the results of a two-stage classification. In the first stage, we classify extreme performers versus the control group. In the second stage, we use the Loser versus Winner model to separate the predicted extreme performers into probable winners and probable losers under a 1:1 relative cost assumption. These results show that firms classified as probable losers consistently underperform the firms classified as probable winners. The effect is most pronounced over the next quarter, but is quite robust across the different relative cost assumptions and statistically significant over the next 12 months. In brief, these strategies return 5.2 to 9.7 percent over the next 6 months, and 9 to 16 percent one-year ahead.

4.5. Supplemental Analyses

Table 7 examines the incremental usefulness of the models in returns prediction after controlling for other known market pricing anomalies. One concern is that the predictive power of the model derives solely from trading strategies already in the literature. This table reports results when future returns are regressed on various investment signals, as well as the output from our predictive models.

To construct this table, we conduct annual regressions of one-year-ahead size-adjusted returns and report the time-series mean and t-statistics for the estimated coefficients.[5] In Panel A, the sample includes all firms in the holdout sample during the 1978–1997 period. In Panel B, the sample consists of only firms classified as "Extremes" by the first-stage model. The independent variables include two indicator variables for our predicted winners and predicted losers (Panel A), a hedge indicator that equals one for winners and zero otherwise (Panel B). We also include five control variables that capture a firm's total accruals to total assets (ACCRUAL), book-to-price (B/P), prior six-month size-adjusted returns (FRTN6), last quarter's earnings surprise (CHGEPS), and firm size (SIZE). To facilitate comparison across the variables, we standardize each control variable by subtracting its cross-sectional mean and dividing by its cross-sectional standard deviation at each portfolio formation date.

Table 7 shows that all five control variables are correlated with future returns in the direction reported in prior studies. The two strongest predictive variables are total accruals (ACCRUAL) and prior returns (FRTN6). Consistent with Sloan (1996), firms with income-increasing accruals earn lower subsequent returns. Also, consistent with Jegadeesh and Titman (1993), firms with higher returns over the past six months earn higher subsequent returns. Both variables have strong predictive power in the total holdout sample (Panel A) as well as among the set of predicted extreme performers (Panel B). In addition, high B/P firms, firms with positive earnings surprises, and smaller firms, also earn higher future returns in the holdout sample (Panel A). B/P continues to predict returns in the smaller sample of predicted extreme performers, but SIZE and CHGEPS have little incremental predictive power (Panel B).

Of greater interest is the predictive power of the signals generated by the models, after controlling for these other five known pricing anomalies. Panel A shows that in the larger sample of all holdout firms, predicted extreme losers earn lower returns than the other firms. The strongest result is obtained under a relative cost assumption of 10 : 1. Firms predicted

Table 7. Correlation with future returns after controlling for other factors.

This table reports results when future returns are regressed on various investment signals. To construct this table, we conduct annual regressions of one-year-ahead size-adjusted returns and report the time-series mean and t-statistics for the estimated coefficients. In Panel A, the sample includes all firms in the holdout sample during the 1978–1997 period. We require at least 10 predicted extreme observations per year. In Panel B, the sample consists of only firms classified as "Extremes" by the first-stage model. The independent variables include two indicator variables for our predicted winners and predicted losers (Panel A), a hedge indicator that equals one for winners and zero otherwise (Panel B), and five control variables. These variables capture a firm's total accruals to total assets (ACCRUAL), book-to-price (B/P), prior six month size-adjusted returns (FRTN6), last quarter's earnings surprise (CHGEPS), and firm size (SIZE). To facilitate comparison, we standardize each control variable by subtracting its cross-sectional mean and dividing by its cross-sectional standard deviation at each portfolio formation date.

Panel A: Annual regressions using all firms in holdout sample (29,794 firm-quarters)

Relative Error Cost Assumption	Constant	Extreme Winners	Extreme Losers	ACCRUAL	B_P	FRTN6	CHGEPS	SIZE
10 to 1	.0062	−.0345	−.1307	−.0278	.0065	.0350	.0064	−.0111
	(2.11)	(−1.57)	(−3.17)	(−9.56)	(2.06)	(11.14)	(2.03)	(−3.57)
20 to 1	.0141	.0005	−.0754	−.0240	.0054	.0311	.0031	−.0014
	(3.69)	(.06)	(−8.40)	(−8.17)	(1.72)	(9.86)	(.98)	(−3.87)
30 to 1	.0092	.0210	−.0384	−.0240	.0051	.0295	.0016	−.0065
	(1.69)	(2.04)	(−4.93)	(−8.11)	(1.64)	(9.32)	(.51)	(−1.57)
40 to 1	.0026	.0301	−.0227	−.0244	.0054	.0294	.0019	−.0026
	(.42)	(2.75)	(−2.85)	(−8.24)	(1.73)	(9.25)	(.60)	(−.61)

Panel B: Annual regressions using only firms predicted to be extreme performers

Relative Error Cost Assumption	Constant	Hedge	ACCRUAL	B/P	FRTN6	CHGEPS	SIZE	Total *N*
10 to 1	−.0781	.0925	−.0436	.0067	.0165	−.0138	−.0507	1,892
	(−1.01)	(1.06)	(−1.78)	(.21)	(.59)	(−.56)	(−1.72)	
20 to 1	−.0611	.0963	−.0261	.0110	.0355	−.0010	−.0185	11,057
	(−5.06)	(5.27)	(−3.62)	(1.45)	(4.51)	(−.12)	(− .22)	
30 to 1	−.0289	.0639	−.0289	.0092	.0367	.0018	−.0023	11,487
	(−4.12)	(5.23)	(−6.06)	(1.82)	(7.10)	(.36)	(−.40)	
40 to 1	−.0202	.0543	−.0289	.0091	.0365	.0015	−.0019	20,654
	(−3.32)	(4.85)	(−6.74)	(1.99)	(7.85)	(.33)	(−.36)	

to be extreme losers under this assumption earned −13.1% lower returns than other firms over the next 12 months, even after controlling for the other variables. The effect is weakest under a 40 : 1 cost assumption, whereby Extreme Losers under-perform by −2.3% over the next 12 months. Results for the predicted extreme winners are generally weaker. The strongest result is an out-performance of 3.0%, based on a 40:1 cost assumption. Overall, it appears that the model has some incremental predictive power for out-of-sample firms even after controlling for five key predictive variables from prior studies.

Panel B focuses on the predictive power of the second-stage regression among firms that were predicted to be extreme performers by the first-stage model. The Hedge variable in these regressions assumes the value one for predicted winners and zero otherwise. This panel shows that predicted winners outperform predicted losers under all four relative cost assumptions. The magnitude of the out-performance ranges from 9.6% (20:1 cost assumption) to 5.4% (40 : 1 cost assumption). These results show that the second-stage model is incrementally useful in predicting out-of-sample returns even after controlling for the other predictive variables.

As a final test, we compare the abnormal returns around subsequent earnings announcement dates for firms classified as predicted extreme winners and predicted extreme losers. Table 8 reports the mean size-adjusted returns accumulated over the three-day period ($t-1$ to $t+1$) surrounding each of the next four quarterly earnings announcements in the year after portfolio formation. All firm-quarters in the holdout sample over the 1977–1997 period are included. The last two columns report the aggregate size-adjusted return for all four announcements, as well as the mean size-adjusted return for the 12 months after portfolio formation.

Table 8 shows that the abnormal return around earnings announcements is consistently lower for predicted losers. In all four quarters, and under all four cost assumptions, the predicted winners earn higher returns than predicted losers. The effect is strongest in quarter $t+1$, and tapers off over the next three quarters. In aggregate, the abnormal returns accumulated over the 12 trading days surrounding the next four earnings releases (2.8% to 5.8%) represent approximately one-third of the total abnormal returns earned by the trading strategy (8.7% to 17.8%). Unless firms' fundamental risk experiences predictable directional shifts around subsequent earnings release dates, these abnormal returns are difficult to reconcile with a risk-based explanation. Rather, they suggest that at least a portion of the abnormal returns is related to market misperceptions about firms' future earnings.

Table 8. Abnormal returns around subsequent earnings announcements for firms classified as extreme winners and extreme losers.

This table reports mean size-adjusted returns accumulated over the three-day period ($t-1$, t, $t+1$) surrounding each of the next four quarterly earnings announcements in the year after portfolio formation. All firm-quarters in the holdout sample over the 1977–1997 period are included. This table also shows the aggregate size-adjusted return for the four earnings announcements, as well as the mean size-adjusted return for the 12 months after portfolio formation. We report results separately for firms classified as extreme winners, extreme losers, and the difference between these two portfolios. * (**) [***] indicates rejection of the null that mean returns for classified winners and classified losers are equal at the 10%, 5%, and 1% levels based on two-tailed tests.

Relative Cost Assumption	Classified Losers?	N	Quarter 1	Quarter 2	Quarter 3	Quarter 4	All Quarters	12-Month Returns
10 to 1	Yes	429	0.002	−0.004	−0.006	0.003	−0.004	−0.099
10 to 1	No	876	0.045	0.008	−0.001	0.001	0.054	0.079
10 to 1	*Difference*		*0.043* ***	*0.012* *	*0.005*	*−0.002*	*0.058* ***	*0.178* ***
20 to 1	Yes	3,961	−0.003	0.001	0.001	0.001	−0.003	−0.050
20 to 1	No	4,300	0.020	0.007	0.003	0.003	0.032	0.062
20 to 1	*Difference*		*0.023* ***	*0.006* ***	*0.003*	*0.003*	*0.035* ***	*0.112* ***
30 to 1	Yes	8,132	−0.002	0.001	0.001	0.001	0.001	−0.031
30 to 1	No	5,974	0.017	0.007	0.003	0.003	0.029	0.067
30 to 1	*Difference*		*0.019* ***	*0.006* ***	*0.002* *	*0.002* *	*0.028* ***	*0.097* ***
40 to 1	Yes	9,727	−0.001	0.001	0.001	0.001	0.002	−0.025
40 to 1	No	6,305	0.017	0.007	0.003	0.003	0.030	0.066
40 to 1	*Difference*		*0.018* ***	*0.006* ***	*0.003* **	*0.002* *	*0.028* ***	*0.087* ***

5. Conclusion

This study examines the predictability of extreme price performance using contextual fundamental analysis. Our results show that a combination of market-related and accounting-based variables have significant predictive power to discern likely extreme price performers 4 to 6 months in advance of the price move. While market-related variables are more useful in identifying extreme performers relative to the general group of control firms, we find that accounting-based fundamental signals are more useful in separating potential extreme winners from extreme losers.

An interesting feature of our research design is the use of a two-stage prediction model. In the first stage, we identify firms that are likely to be extreme winners or extreme losers relative to a control sample. This provides the context for our fundamental analysis. We show that extreme losers and extreme winners share many common traits. These common traits make it difficult to isolate "torpedo" stocks (extreme losers) from "rocket" stocks (extreme winners) using standard techniques. Our solution is to first identify the extreme performers, then estimate a second model that distinguishes winners from losers, conditional on the fact that they are both extreme performers.

We find that this two-stage approach can identify a substantial proportion of the extreme losers and winners. In the first stage, we show that extreme performers tend to be younger (AGE), smaller (SIZE) firms, with higher recent trading volume (AVGVOL), higher sales growth (SGI), greater returns volatility (STDRET), higher R&D intensity, and lower sales-to-price (S/P) ratios than the control group. The portrait that emerges is that of a set of young growth firms engaged in more speculative ventures. In this first-stage, accounting variables played a relative minor role.

In the second stage, we find that extreme losers are growth firms with weaker financial performance that appear in a number of dimensions. Leading indicators of their upcoming woes include lower sales growth (SGI), deteriorating margins (GMG), lower R&D spending (R&D), more negative earnings surprises (CHGEPS), worse recent price performance (FRTN6), more aggressive accruals (ACCRUAL), and higher capital expenditures (CAPX). In this second-stage regression, accounting variables play a primary role in separating predicted winners from predicted losers.

Finally, we show that in out-of-sample tests, firms classified as winners earn much higher subsequent returns than firms classified as losers. Our first-stage results show that the mean returns for the predicted extreme firms are not distinguishable from the mean returns for the predicted control groups. However, our second-stage results show that predicted winners outperform predicted losers by 8.7% to 17.8% over the next 12-months, depending on the relative cost assumption. These results are virtually unaffected by the inclusion of five control variables: accruals, B/P, price momentum, earnings surprise, and firm size. A significant portion of the future returns are earned around the next four earnings announcements, suggesting they are related to market misperceptions about future earnings.

These results illustrate the usefulness of contextual fundamental analysis in growth investment strategies. While market-based variables are useful in identifying potential extreme price movements, it is accounting-based fundamental signals that are most useful in separating the winners from the losers. Our findings are strikingly complementary to a recent study by Piotroski (2000), which focuses on the set of value (high B/P) stocks. That study shows that fundamental accounting signals are useful in separating eventual winners from losers among high B/P firms. Our sample consists of growth firms rather than value firms. However, we also find fundamental signals important in separating winners from losers. In both studies, the incremental power of accounting variables is enhanced by first identifying a subset of firms with similar primary attributes.

Taken together, the results in these two papers illustrate the contrast between value and growth investing. Most of Piotroski's variables are, appropriately, nominated from the literature on financial distress. In contrast, a number of our variables focus on growth management and growth opportunities. Some variables identified by Piotroski (2000) also have predictive power for our growth-oriented stocks—i.e., total accruals (ACCRUAL), price momentum (FRTN6), gross margin (GMG), and trading volume (AVGVOL). However, we have identified a number of additional variables that seem particularly important in the growth context—SGI, CHGEPS, CAPX, and R&D. These variables are growth and cash flow related, and seem to provide early indications of changes in the growth trajectory for "glamour" firms.

Looking ahead, it seems likely that users of financial information can enhance the power of accounting data by identifying other fruitful partitions among the total population of firms. Financial analysts generally conduct their analysis by focusing on subsets of firms that have common attributes. Perhaps future research in fundamental analysis would be more fruitful if academics follow a similar strategy.

Acknowledgment

This paper has improved greatly from the comments and suggestions of Tom Dyckman, Kenneth Levy, and Richard Sloan (Editor), as well as workshop participants at Cornell, Dartmouth, Indiana, Michigan State, Barclays Global Investors, the Prudential Securities Quantitative Research Conference, and the *Review of Accounting Studies* (2000) Conference. Information on analyst forecasts is provided by I/B/E/S Inc.

Notes

1. Practitioners are concerned with avoiding "torpedo" stocks and picking "rockets" (Skinner and Sloan, 1998). In the same spirit, Harvey and Siddique (2000) suggest investor preferences extend beyond mean and variance of expected returns, to include skewness measures (e.g., see Harvey and Siddique, 2000).
2. Prior studies that examine the predictability of returns include Basu (1977), Davis (1994), Haugen and Baker (1996). Jegadeesh and Titman (1993), Lee and Swaminathan (2000), Sloan (1996), and Bernard and Thomas (1989, 1990). We discuss these and other related studies in more detail in the next section.
3. The literature on accounting fraud detection suggests that the greater capital needs of young firms tend to put pressure on managers to achieve earnings targets and maintain the appearance of growth (National Commission on Fraudulent Financial Reporting (1987), Fridson (1996), Beneish (1997)). These firms also tend to have weaker governance structures and internal controls (Loebbecke et al., 1989; Beasley, 1996), less developed information environments (Barry and Brown, 1984), and are at greater risk of financial distress (Dopuch et al. 1987).
4. Asness (1997) shows that value and growth strategies are negatively correlated. It is therefore perhaps not surprising to find that value multiples are not particularly effective in forecasting returns among extremely volatile growth stocks.
5. We have also conducted quarterly regressions with quarterly holding periods. The results are very similar.

References

Abarbanell, J. S. and B. J. Bushee. (1998). "Abnormal Returns to a Fundamental Analysis Strategy." *The Accounting Review* 73, 19–45.

Altman, E. I., R. G. Haldeman and P. Narayanan. (1977). "Zeta Analysis: A New Model to Identify Bankruptcy Risk of Corporations." *Journal of Banking and Finance* 1, 29–54.

Asness, C. A. (1997). "The Interaction of Value and Momentum Strategies." *Financial Analysts Journal* 53, 29–36.

BARRA, *United States Equity Model Version 3 (E3), Model Handbook*, BARRA Corporation: Berkeley, CA.

Barry, C. B. and S. J. Brown. (1984). "Differential Information and the Small Firm Effect." *Journal of Financial Economics* 13, 283–294.

Basu, S. (1977). "Investment Performance of Common Stocks in Relation to Their Price Earnings Ratios: A Test of the Efficient Market Hypothesis." *Journal of Finance* 32, 663–682.

Beasley, M. S. (1996). "An Empirical Analysis of the Relation Between Board of Director Composition and Financial Statement Fraud." *The Accounting Review* 71, 443–465.

Beaver, W., P. Kettler and M. Scholes. (1970). "The Association Between Market Determined and Accounting Determined Risk Measures." *The Accounting Review* 45, 654–682.

Beneish, M. D. (1997). "Detecting GAAP Violation: Implications for Assessing Earnings Management Among Firms with Extreme Financial Performance." *Journal of Accounting and Public Policy* 16, 271–309.

—, (1999a). "The Detection of Earnings Manipulation." *Financial Analysts Journal* 55, 24–36.

—, (1999b). "Incentives and Penalties Related to Earnings Overstatements that Violate GAAP." *The Accounting Review* 74, 425–457.

Beneish, M. D. and C. R. Harvey. (1998). "Measurement Error and Nonlinearity in the Earnings-Returns Relation." *Review of Quantitative Finance and Accounting* 3, 219–247.

Bernard, V. and J. Thomas. (1989). "Post-Earnings-Announcement Drift: Delayed Price Response or Risk Premium?" *Journal of Accounting Research* 27, 1–48.

Bernard, V. and J. Thomas. (1990). "Evidence that Stock Prices do not Fully Reflect the Implications of Current Earnings for Future Earnings." *Journal of Accounting & Economics* 13, 305–340

Chan, Louis K. C., Y. Hamao and J. Lakonishok. (1991). "Fundamentals and Stock Returns in Japan." *Journal of Finance* 46, 1739–1764.

Chan, L. K., J. Lakonishok and T. Sougiannis. (1999). "The Stock Market Valuation of Research and Development Expenditures." Working Paper, University of Illinois.

Christie, A. A. (1982). "The Stochastic Behavior of Common Stock Variances: Value, Leverage and Interest Rate Effects." *Journal of Financial Economics* 10, 407–432.

Davis, J. L. (1994). "The Cross-Section of Realized Stock Returns: The Pre-Compustat Evidence." *Journal of Finance* 49, 1579–1593.

Dietrich, R. J. and R. S. Kaplan. (1982). "Empirical Analysis of the Commercial Loan Classification Decision." *The Accounting Review* 57, 18–38.

Dopuch, N., R. W. Holthausen and R. W. Leftwich. (1987). "Predicting Audit Qualifications with Financial and Market Variables." *The Accounting Review* 62, 431–454.

Fama, E. F. and K. R. French. (1992). "The Cross-Section of Expected Stock Returns." *Journal of Finance* 47, 427–465.

Fridson, M. S. (1996). "Financial Statement Analysis: A Practitioner's Guide." New York: John Wiley.

Harvey, C. R. and A. Siddique. (2000). "Conditional Skewness in Asset Pricing Tests." *Journal of Finance* 55, 1263–1295.

Haugen, R. A. and N. L. Baker. (1996). "Commonality in the Determinants of Expected Stock Returns." *Journal of Financial Economics* 41, 401–439.

Hayn, C. (1995). "The Information Content of Losses." *Journal of Accounting and Economics* 20, 125–153.

Holthausen, R. W. and D. F. Larcker. (1992). "The Prediction of Stock Returns Using Financial Statement Information." *Journal of Accounting and Economics* 15, 373–411.

Jegadeesh, N. and S. Titman. (1993). "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency." *Journal of Finance* 48, 65–91.

Lakonishok, J., A. Shleifer and R. W. Vishny. (1994). "Contrarian Investment, Extrapolation, and Risk." *Journal of Finance* 49, 1541–1578.

LaPorta, R. (1996). "Expectations and the Cross-Section of Stock Returns." *Journal of Finance* 51, 1715–1742.

Lee, C. M. C. and B. Swaminathan. (2000). "Price Momentum and Trading Volume." *Journal of Finance* 55, 2017–2070.

Lev, B. and R. Thiagarajan. (1993). “Fundamental Information Analysis.” *Journal of Accounting Research* 31, 190–202.

National Commission on Fraudulent Financial Reporting, (1987). *Report of National Commission on Fraudulent Financial Reporting.* New York: AICPA.

O’Shaughnessy, J. P. (1998). *What Works on Wall Street,* Revised Edition, McGraw-Hill, New York.

Ou, J. A. and S. H. Penman. (1989). “Financial Statement Analysis and the Prediction of Stock Returns.” *Journal of Accounting and Economics* 11, 295–329.

Palepu, K. G. (1986). “Predicting Takeover Targets: A Methodological and Empirical Analysis.” *Journal of Accounting and Economics* 27, 3–36.

Piotroski, J. (2000). “Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers.” *Journal of Accounting Research* (forthcoming).

Reinganum, M. R. (1998). “The Anatomy of a Stock Market Winner.” *Financial Analysts Journal,* March/April, 16–28.

Seyhun, H. N. and M. Bradley. (1997). “Corporate Bankruptcy and Insider Trading.” *Journal of Business* 70, 189–216.

Skinner, D. J. and R. G. Sloan. (1998). “Earnings Surprises, Growth Expectations and Stock Returns, or, don’t let a Torpedo Stock Sink Your Portfolio.” Working Paper, University of Michigan July.

Sloan, R. G. (1996). “Do Stock Prices Fully Reflect Information in Accruals and Cash Flows About Future Earnings?” *The Accounting Review* 71, 289–315.

Zmijewski, M. E. (1984). “Methodological Issues Related to the Estimation of Financial Distress Prediction Models.” *Journal of Accounting Research* 22, 59–82.