On Distinguishing Between Valuation and Arbitrage Motivated Short Selling

Hemang Desai
Edwin L. Cox School of Business
Southern Methodist University
Dallas, TX 75275-0333
214 768-3185
hdesai@mail.cox.smu.edu

Srinivasan Krishnamurthy
School of Management
SUNY – Binghamton University
Binghamton, NY 13902
607 777-6861
srinik@binghamton.edu

and

Kumar Venkataraman
Edwin L. Cox School of Business
Southern Methodist University
Dallas, TX 75275-0333
214 768-7005
kumar@mail.cox.smu.edu

First Draft: March 2006
This Draft: September 2006

We thank Hank Bessembinder, Alex Butler, Nat Guild, Ryan LaFond, Michael Long, Kristian Rydqvist, Ajai Singh, Mary Stanford, Rex Thompson, Brett Trueman, and seminar participants at UCLA, Duke, Penn State, SUNY at Binghamton, and University of Texas at Dallas for helpful discussions and comments; Honghui Chen, Muku Santhanakrishnan and Vijay Singal for providing some of the data used in the study; and Machiko Hollifield, Harish Raman and Jagadish Pasunoori for research assistance.
On Distinguishing Between Valuation and Arbitrage Motivated Short Selling

Abstract

The short interest data reported in the United States aggregate valuation shorts (motivated by a pessimistic opinion on firm value) and arbitrage shorts (motivated by various arbitrage or hedging strategies). However, the information content of these two sources of short interest is different and hence their association with future returns is expected to be different. In recent years, the association between short interest and future returns has weakened considerably, reflecting the increasing importance of institutions that execute arbitrage strategies. The primary contribution of this study is an empirical model that ex-ante helps differentiate valuation shorts from arbitrage shorts. In out-of-sample tests, we document that, consistent with theoretical predictions, the firms identified by the model as valuation shorts exhibit high short interest and poor future returns. Furthermore, firms identified as arbitrage shorts do not exhibit significant negative returns but instead exhibit characteristics associated with arbitrage strategies. The paper also identifies variables that are correlated with the information set of short sellers. We present an application that exploits the model’s ability to ex-ante identify poor performers and discuss broader applications in other finance settings.

Keywords: Informed short selling, fundamental analysis, stock returns.
Short selling occurs when investors sell shares that they do not own, borrowing the shares from an equity lender and delivering them to the buyer. Short sellers expect to repay the stock loan by purchasing shares from the open market at a lower price at a later date. However, various legal and institutional constraints make short selling a costly endeavor. Miller (1977) postulates that such constraints may dissuade relatively pessimistic investors from selling short. On the other hand, such constraints do not deter an optimist from going long, causing the stock to be overpriced, especially when beliefs are widely dispersed. Using a rational expectations framework, Diamond and Verrecchia (1987) argue that only informed investors with very negative information will typically sell short, as short selling is quite costly. Collectively, the literature predicts that short sales convey negative information about firm value. Using the level of short interest as a proxy for the adverse information held by short sellers, prior empirical evidence suggests that short sellers are informed investors and that short interest is a bearish indicator for firm value.¹

The short interest data reported in the United States aggregate both short selling that is driven by a pessimistic opinion on firm valuation (valuation shorts) and by various arbitrage or hedging strategies (arbitrage shorts). While valuation shorts are motivated by a bearish view on the stock, arbitrage shorts convey little information about the perceived overvaluation of the stock, as their objective is to exploit relative mispricing.² Thus, the information content of the two sources of short selling is quite different. However, prior research had to rely on short interest data to proxy for the adverse information held by short sellers.


² Examples of hedging or arbitrage related short selling include short positions undertaken by convertible arbitrageurs who typically long the convertible bond and short the underlying stock, or by merger arbitrageurs who long the target and short the acquirer in stock for stock mergers, or by index arbitrageurs who assume long-short positions based on pricing discrepancies between the index futures and the underlying stocks. Mitchell, Pulvino, and Stafford (2004) conclude that, in the case of merger arbitrage and convertibles arbitrage, short sales due to such arbitrage activities is likely to be much larger than information or valuation based short sales.
interest as a proxy for bearish sentiment in the stock, mainly due to data limitations. Disentangling the components of short interest was not as critical for past research because the incidence of arbitrage shorts was relatively modest. However, in recent years, short sales motivated by arbitrage strategies has increased significantly, reflecting the explosive growth in hedge funds and institutions that engage in various “market neutral” investment strategies. Confirming this trend, Asquith, Pathak and Ritter (2005) document that the percentage of shares held short has increased over 1988 to 2002. More importantly, they show that the association between the level of short interest and future returns has weakened in recent years. This suggests that the level of short interest is an imprecise proxy for the adverse information held by short sellers, especially in recent years. Given the continuing growth in institutions that engage in arbitrage strategies, this problem is expected to get more acute over time, suggesting that distinguishing between valuation shorts and arbitrage shorts presents an important avenue for future research. The primary contribution of this study is an empirical model that ex-ante helps differentiate valuation shorts from arbitrage shorts. Such a model would increase the statistical power of empirical tests that rely on the level of short interest to proxy for the adverse information held by short sellers.

We also present new evidence on the information set of short sellers. Although prior theoretical and empirical research has characterized short sellers as informed investors, the empirical evidence on their information set is relatively sparse. We document that short sellers rely on information contained in valuation multiples, such as book-to-market ratio, and financial statement ratios, such as accruals and sales growth, to identify target firms. This is an important result because recent empirical work supports the notion that short sellers facilitate the flow of

---

3 During the time period covered by our study (1990-2004), the total number of hedge funds has increased from 610 to 7,436 and the assets under management from USD 39 billion to USD 972 billion (see Agarwal and Naik (2005)).
private information into stock prices (see Chen and Singal (2003), and Bris, Goetzmann and Zhu (2003)). Our findings suggest that informed arbitrageurs exploit the information in variables (such as earnings quality and valuation ratios) that prior work has shown to be related to future returns in identifying target firms. These findings present new evidence on the information set and trading behavior of an important group of information arbitragers, thereby furthering our understanding of the information arbitrage process.

We utilize a unique database of firms that were identified by an independent research firm as targets for short selling. The short database includes every report issued by the firm since its inception in 1998 until June 2005. These short recommendations are motivated by perceived overvaluation of the firm and are unrelated to hedging or arbitrage related strategies. We document that these firms experience a raw (market-adjusted) return of -3.8% (-4.5%) in the month when the short report was issued and –21.9% (-28.0%) in the subsequent 12 months. The poor performance of firms in the short database is consistent with the notion that the research firm is successful in identifying poor performers. Using the short database, we build a parsimonious model to identify short selling motivated by a pessimistic opinion on firm valuation. We estimate a logistic model over the years 1997-2004, where the dependent variable equals ‘1’ for firms in the short database, and equals ‘0’ otherwise. The results indicate that, when identifying short targets, the research firm is sensitive to information contained in valuation multiples, fundamental variables, and indicators of earnings quality.

We conduct several tests during the out-of-sample period 1990-1996 and document that the model can forecast both the level of short interest and abnormal returns. Specifically, we sort firms into decile portfolios constructed each year based on the predicted probability from the model. We document a monotonic trend in short interest across decile portfolios in the out-of-
sample period, increasing from about 0.5% for firms in the lowest decile to over 3.2% for firms in the highest decile. A monotonic pattern is also observed for average monthly abnormal returns. The intercept from a regression of monthly portfolio returns on the three Fama-French factors decreases from 1.25% for firms in the lowest decile to about –0.76% for firms in the highest decile. These abnormal returns are both economically large and statistically significant. Thus, although the short interest model is developed from a small set of potential short targets identified by one firm, the model is able to forecast short interest and returns for a broader sample in an out-of-sample period, implying that the selection criteria appear to be correlated across short sellers and, to some extent, constant across time.

Further tests suggest that the short interest model can successfully identify valuation and arbitrage shorts. We independently categorize stocks into decile portfolios based on the (observed) short interest ratio and into three groups based on the predicted probability from the model (30%, 40%, and 30%). We focus on the firms in the highest and the lowest short interest deciles and examine the subsequent abnormal returns for firms in each of the six groups. Intuitively, this categorization identifies firms with high short interest and high predicted probability as valuation shorts and firms with high short interest but low predicted probability as arbitrage or hedging shorts. Consistent with this categorization, the average monthly abnormal return for firms identified as valuation shorts is negative and significant. In contrast, the abnormal return for firms identified as arbitrage shorts is not statistically significant. Firms identified by the model as arbitrage shorts have high book-to-market ratios (value firms), are more likely to have convertible bonds outstanding, and are more likely to be included in the S&P 500 index, suggesting that short selling in these firms is motivated by arbitrage reasons. On the other hand, firms identified by the model as valuation shorts exhibit low book-to-market ratios.

---

4 The results are robust to the use of a four-factor model that includes a momentum factor.
(glamour firms), are less likely to have convertible bonds outstanding, and are less likely to be included in the S&P 500 index. Interestingly, firms with high predicted probability but low short interest also exhibit significantly negative abnormal returns. However, these firms are small and have poor liquidity, suggesting that short sellers avoid these firms due to significant impediments to assuming short positions.

The short interest model can also help design better tests of competing theories in many applications, as the model identifies firms that are likely to perform poorly. Presumably, ex-ante sorting sample firms into subsets that display similar characteristics, as suggested by theory, would enhance the power of the statistical tests. As an application, we show that the model can distinguish high momentum firms that experience poor subsequent returns from those that do not. This allows us to classify high momentum stocks into relatively homogenous subsets that are either prone to a reversal (as suggested by Hong and Stein (1999)), or where prior patterns are likely to persist. We also suggest other settings where such an approach would be valuable.

The rest of the paper proceeds as follows. Section I briefly discusses related literature and the hypotheses. Details about the data, summary statistics, and methodology are provided in Section II. Section III presents the logistic model and describes the out-of-sample findings and robustness tests. The application pertaining to return momentum is described in Section IV. The conclusions are presented in Section V.

I. Prior Literature and Hypotheses

In this section, we briefly summarize the related literature and develop the hypotheses that underlie the empirical tests that follow. The first strand of literature analyzes whether short interest is informative about a firm’s future performance, and examines various characteristics
that are associated with the decision to short a stock. The second line of research attempts to separate valuation and hedging shorts, and is still at a nascent stage.

A. Performance and Firm Characteristics Associated with Short Selling


Researchers have also studied the relation between short interest and firm characteristics such as valuation ratios, earnings quality indicators, and liquidity. Dechow, Hutton, Meulbroek and Sloan (2001) document that short sellers target firms with high multiples of price relative to fundamentals such as earnings and book value. Desai, Krishnamurthy and Venkataraman (2006) find that short sellers target firms that subsequently restate their earnings and, in particular, target restating firms with high accruals. Their findings suggest that short sellers are sensitive to earnings quality indicators. Hishleifer, Teoh and Yu (2005) show that short sellers assume positions based on various accounting based anomalies. D’Avolio (2002) finds that short sellers have difficulty in establishing positions in smaller, illiquid securities and hence prefer somewhat
larger and liquid stocks. Our choice of explanatory variables in the short interest model is intended to capture the information attributes described above in a parsimonious manner.

B. Valuation-based versus Hedging-based Short Sales

As discussed earlier, short sales could either be motivated by a bearish view of the stock or they could be motivated by certain hedging or arbitrage activities that are not intended to convey negative information about the firm. In an attempt to discriminate between valuation shorts and arbitrage shorts, Asquith et al. (2005) classify high short interest firms with convertible bonds as arbitrage shorts and the remaining high short interest firms as valuation shorts. During the time period covered by our study (1990-2004), firms with convertible bonds outstanding comprised only about 20% of firm-years (excluding financials and firms in regulated industries) on Compustat. Clearly, many firms without convertible bonds outstanding could also be the target of arbitrage shorts. Asquith et al. (2005) acknowledge that classifying firms as arbitrage shorts based on convertible bonds alone is imprecise, suggesting that distinguishing valuation and arbitrage shorts presents an important avenue for future research.

If the short interest model sufficiently describes the information set of valuation-based short sellers, we expect that the model should forecast both the level of short interest and abnormal returns in an out-of-sample period. Further, if the model can distinguish between valuation and arbitrage shorts, we expect that firms identified as valuation shorts should experience poor subsequent returns. In contrast, firms identified as arbitrage shorts should not exhibit negative returns but instead should exhibit characteristics associated with arbitrage strategies, such as index membership and convertible bonds outstanding. The following hypotheses summarize these arguments:

\textit{H1: The short interest model should forecast the level of short interest and abnormal returns.}
**H2: Controlling for short interest, which aggregates both valuation and arbitrage shorts, we expect that (a) firms identified as valuation shorts should experience negative abnormal returns, and (b) firms identified as arbitrage shorts should exhibit characteristics that are associated with arbitrage strategies.**

**II. Data and Methodology**

**A. Sample Selection**

We obtain the data for the study from an independent research firm. At periodic intervals, the firm alerts its subscribing clients about potential short targets via a detailed research report. The first report was issued in September 1998 and the last report available to us was issued in June 2005. The firm issues about 8-10 reports a year. While the firm assimilates information from many sources, we were informed that they primarily rely on their own analysis and avoid conventional Wall Street sources such as brokerage analyst reports, conference calls, discussions with corporate executives, etc. The primary objective of the report is to identify a promising short target and to present arguments detailing why the prior performance of the identified firm is not sustainable and is likely to reverse significantly. Thus, the sample targets are clearly identified as valuation shorts and are not hedging shorts. Each report begins with a brief history and description of the firm and its business, followed by an analysis of the firm’s financials, the firm’s and the industry’s growth potential, and the competitive environment in which the firm operates. Often the report will question the use of aggressive accounting practices. The report also tracks insider sales and at times questions the firm’s governance.

The short database consists of 67 firms identified in the reports during 1998-2005 as promising short candidates. We exclude (a) two firms that appeared more than once in the database, retaining only the first occurrence, (b) seven firms in the Financials (SIC 6000-6999), Utilities (SIC 4900-4999) or Communications (SIC 4800-4899) industries, and (c) four firms
with insufficient data on CRSP and Compustat in the year prior to being listed in the short database. After these screens, the short database contains 54 firms.5

In selecting control firms, we retain all firms with Compustat data since 1990 and exclude firms (a) with sales or total assets less than or equal to $1 million, (b) with a positive ADR ratio (Compustat annual data item #234), and (c) in the financial, communications, and utilities industry, thus yielding 87,716 firm-year observations. We retain firm-years if CRSP data is available, thus eliminating Compustat observations that relate to non-public firms and subsidiaries. The CRSP screen reduces the sample to 66,022 firm-year observations.

Since the short database covers the period 1998-2005, we classify the eight-year period 1997-2004 as the estimation period (35,143 observations) and the seven-year period 1990-1996 as the out of sample prediction period (30,879). In the estimation period, we retain one firm-year for each sample firm, corresponding to the fiscal year preceding the date when the firm was identified in the short database, thus eliminating 295 observations. The remaining firm-year observations are classified as control firms. Thus, the estimation period data consists of 54 firm-years for sample firms and 34,794 firm-years for control firms. For the out of sample prediction period, a similar selection procedure leaves us with 30,879 firm-year observations. Since not all firm-years will have the necessary data for all variables used in the regression analysis, the actual number of observations will vary depending upon data availability.

B. Model

The short interest model is estimated using 54 firm-years as sample observations and 34,794 firm-years as control observations. Our approach takes annual snapshots of the financial information for the cohort of sample and control firms every September during the estimation

---

5 Some variables, such as operating and total accruals, cannot be computed meaningfully for financial firms. Firms in two-digit SIC codes 48 and 49 were excluded as these represent regulated industries.
period. The following example illustrates the procedure for matching firm-years with the financial data. Consider all firms with Compustat fiscal year of 1998. Given Compustat’s reporting convention, the fiscal year-end for all these firms would fall between June 1998 and May 1999. Assuming a four month reporting lag, the data for fiscal year 1998 would be available for all cohort firms (year = 1998) by September 1999. Thus, we use the 1998 fiscal year-end financial data for the annual snapshot of firms in September 1999. Further, for sample firms identified as short targets during the period from October 1999 to September 2000, the pre-event financial data is from the 1998 fiscal year.

During the prediction period, we follow a similar approach and calculate the predicted probabilities every September based on the coefficients from the short interest model and the available annual financial information for all cohort firms by September. Firms are then assigned annually to decile portfolios based on predicted probabilities every September and remain in the assigned decile portfolio from October through September of the next year. The short interest ratio for the portfolio is computed in October and the abnormal returns are estimated over a twelve-month holding period, from October to September of the next year. In other words, for the fiscal year 1995, all the relevant annual data are available for all firms by September 1996. The short interest for these firms is computed in October 1996 and the abnormal returns are estimated from October 1996 to September 1997.

The short interest model can be expressed as follows:

$$S_i = L (\alpha + \beta_i * X_i + C_i)$$  (1)

where $S_i$ is a indicator variable that equals ‘1’ for sample firms identified in the short database.

---

6 This approach ensures that the annual accounting information for all the sample and control firms (universe of eligible Compustat) firms is available when the model is estimated.

7 The fiscal year 1998 is designated as year -1. Since our firms were targeted in calendar years 1998-2005, the Compustat years corresponding to year –1 spans 1997-2004, which we call the estimation period.
and equals ‘0’ for control firms, $X_i$ is the vector of variables selected to proxy for the information set of short sellers, $\varepsilon_i$ is the residual error term, and $L$ indicates that the model is based on a logistic regression.

The selection of the explanatory variables relies on evidence from extant research on short seller behavior and is designed to capture the information set of informed short sellers in a parsimonious manner. Following prior literature (discussed in Section I), we include valuation multiples, financial statement variables that proxy for earnings quality, and firm characteristics as explanatory variables. The valuation indicators are the equity book-to-market ratio $BM$ and prior momentum, measured as one-year buy-and-hold return over the period October$_{t-1}$ to September$_t$. Book value of total assets (SIZE) and average share turnover ($TURNOVER$) capture the short sellers’ reluctance to take positions in small, illiquid stocks.$^8$ $TURNOVER$ is calculated as the arithmetic average of the daily share turnover (ratio of shares traded to total shares outstanding) over the period October$_{t-1}$ to September$_t$.9

We include financial statement variables that are designed to capture the quality of earnings.$^{10}$ We include two measures of accruals, namely total accruals ($TOTACC$) and operating accruals ($OPACC$). In addition to accruals, we include seven financial statement variables that are identified by Beneish (1999) as being related to earnings manipulation. The first variable is the days’ sales in receivables index (DSRI). Since a large increase in receivables could indicate

---

$^8$ While institutional ownership may be a better proxy for loanable supply, we do not have access to institutional ownership data at this time. Hence, we use firm size as an alternative proxy, motivated by prior evidence of a strong correlation between size and institutional ownership (see Sias and Starks (1997)). We have replicated the entire analysis using market value of equity instead of book value of total assets and find similar results. Since market value of equity is more closely correlated (relative to book value of assets) with other explanatory variables such as prior return and the book to market ratio, we report results using the latter measure of firm size.

$^9$ The detailed definition of all the variables is presented in Appendix 1.

$^{10}$ There is no universally accepted definition of earnings quality (see Schipper and Vincent (2003)). However, given that short sellers’ interest lies in identifying firms whose performance is not sustainable, we consider earnings to be of poor quality if they are not likely to be sustained and hence use variables that have been shown to be associated with lack of earnings persistence. (see Sloan (1996) and Richardson, Sloan, Soliman and Tuna (2005)).
revenue inflation or relaxation of credit policy to generate higher sales, we expect that DSRI will be positively related to being the target of a valuation short. We include the gross margin index since improved margins accompanied by high accruals might suggest that the increased margins may not be sustainable. The asset quality index $AQI$ measures the extent of capitalization of assets with uncertain benefits, such as goodwill. It may also be indicative of a firm’s propensity to engage in cost deferral by capitalizing expenses. Thus, we expect a positive association between $AQI$ and short selling. Short sellers may target firms with high sales growth ($SGI$), consistent with the notion that firms may inflate their reported revenues in an attempt to mislead investors about future growth prospects. $DEPI$ measures the depreciation rate and indicates whether the firm has made income-increasing accounting choices and/or increased its estimate of the useful lives of depreciable assets. Such tactics delay reporting an earnings decline, which may be a useful signal for short sellers. An increase in leverage ($LVGI$) suggests that debt covenants are more likely to be binding, generating more incentives for financial statement manipulation. Lev and Thiagarajan (1993) suggest that analysts perceive an increase in sales, general and administrative expenses ($SGAI$) as a negative signal about the future prospects of the firm. Therefore, we expect that firms with higher levels of $LVGI$ and $SGAI$ are more likely to be targeted by informed short sellers.

Since financial statement ratios are likely to vary across industries, we estimate the models using both raw and industry-adjusted values. Specifically, for each explanatory variable, we subtract the industry median (calculated annually, industries based on 2 digit SIC code) from the raw values for each firm to compute the industry-adjusted values.

C. Summary Statistics

The sample firms are not clustered in time. When we divide the eight-year estimation
period into two four-year periods (1998-2001 and 2002-2005), we find that each sub-period contains roughly equal number of observations. Furthermore, the maximum number of observations in a given year is ten (in 2002). We find modest evidence of industry concentration. ‘Business Services’ (SIC two-digit code 73) accounts for 14 observations (26%). In addition, chemicals and allied products (SIC 28) and industrial & commercial machinery and computer equipment (SIC 35) each account for four observations (7% each). All other industries (based on two-digit SIC) have fewer than four firms. Thus, the sample represents a fairly broad cross-section of firms, representing 23 different industries.

Table I presents the summary statistics for sample firms in the fiscal year prior to the date of the short report (year -1). Since mean accounting ratios could be affected by outliers, we focus on the medians, although mean values are also reported for completeness. In the year prior to being identified as a short target, the sample firms have low BM ratios (glamour stocks), consistent with the findings in Dechow et al. (2001). The median BM ratio of 0.24 is smaller than the industry median and the difference is significant at the one-percent level. Also, the sample firms have experienced a large run-up in stock price (momentum firms). The mean raw return in the 12 months preceding the estimation (October through September of year -1) is 38.8%. We also find evidence of a significant run-up in prices in the 12 months leading up to the date of the report. Specifically, the sample firms exhibit market-adjusted return of 67.85% (t-statistic of 4.06) from month -12 to month -1, relative to the month in which the report was issued (results not tabulated). Thus, the research firm seems to target low BM firms that have experienced a significant run-up in stock price. The mean (median) value of total assets is $857 million ($405 million) and the market value of equity is $1,090 million ($691 million), suggesting that the research firm does not target large firms. This is not surprising as the
informational advantage of short sellers is unlikely to persist in large firms. The industry adjusted trading volume and turnover statistics suggest that the sample firms are relatively more liquid than their industry cohorts.

From Table I, we also note that the sample firms exhibit superior accounting performance compared to the industry median. Specifically, the median industry-adjusted ROA is 0.96% and is significant at the ten-percent level, based on the Wilcoxon signed rank test. Further, the median total accruals and operating accruals prior to the report date are significantly higher than the industry median. This finding is consistent with Desai et al. (2006) and Hirshleifer et al. (2005), who show that short sellers target firms with high accruals. The sample firms report higher sales growth and a greater improvement in margins than their cohorts. The industry-adjusted sales growth is positive and the industry-adjusted GMI is negative (improved margins over the previous year), both significant at the one-percent level. The other financial ratios are not significantly different when compared to the industry medians.

An examination of the stock market performance after the issuance of the report indicates that the sample firms experience sharp reversals in their performance.  

In the month in which the report is issued (month 0), the market-adjusted return is -4.48% (t-statistic of -1.91). The performance continues to decline in the 12 month period after the issuance of the report and the sample firms underperform the market, on average, by 27.98% (t-statistic of -3.05). Even the mean raw return of the sample firms is large and negative (-21.88%) over the subsequent 12 months. Thus, the firms experience a significant reversal in their fortunes soon after being identified as a promising short target in the report, suggesting that the short database represents an appropriate sample for modeling information-based short selling.

---

11 These results are not tabulated, but are available upon request.
III. Results

In this section, we present the main results of the study. We first describe our estimation period results, where we fit the observed data for Compustat years 1997-2004 using a logistic regression. To test out-of-sample predictability, we estimate the predicted probability using the model coefficients during the out-of-sample prediction period 1990-1996, and sort firms into predicted probability deciles. We test for differences in the (observed) total short interest and the abnormal stock performance across the decile portfolios (Hypothesis H1). The prediction period is based on prior period data because short interest or return data are not available after the end of the estimation period (2004 onwards). We also describe the various additional tests that we conduct to verify that our model is able to distinguish between valuation shorts and arbitrage shorts (Hypothesis H2).

A. Estimation Period Results

In Table II, we report the results of logistic regressions that model the decision process of the research firm. The dependent variable is an indicator variable that equals ‘1’ if the observation is for a sample firm in year –1 and equals ‘0’ otherwise. In models 1 and 2, the explanatory variables are unadjusted, while in models 3 and 4, the explanatory variables are industry-median adjusted. In models 2 and 4, we replace operating accruals with total accruals. The main findings are similar across all the models, suggesting that the findings are robust to alternative accruals measures and to industry effects. In the interests of brevity, the discussions below focus on model 3, although the coefficients from all four models are reported in Table II.

With regards to valuation indicators, the coefficient on prior momentum is positive and that on BM ratio is negative (both significant at the five percent level), suggesting that the research firm behaves as a contrarian, targeting glamour firms that have experienced a large

---

12 To minimize the effect of outliers, all variables are winsorized at the 0.5% and 99.5% level.
price run-up. Lakonishok, Shleifer and Vishny (1994) attribute poor performance of glamour stocks to naïve extrapolation of past performance. Thus, it appears that the research firm is able to identify stocks whose price has been bid up due to such extrapolation and might be expected to reverse. Turning to financial statement variables, the results suggest that the research firm is sensitive to information conveyed by financial ratios. Specifically, the coefficients on operating accruals (OPACC), sales growth index (SGI), and the sales, general and administrative expenses index (SGAI) are positive and significant at the five-percent level or better. Evidence in Sloan (1996) indicates that the earnings of firms with high accruals are strongly mean reverting, suggesting that high accruals are indicative of poor earnings quality. Thus, it is likely that strong sales and price growth coupled with the large accrual component in the reported earnings may have attracted the attention of the research firm. The positive coefficient on SGAI suggests that the research firm views an increase in SG&A as a bearish indicator, consistent with Lev and Thiagarajan (1993). The coefficient on gross margin index is negative, suggesting that the firm targets firms with an increase in gross margin, ceteris paribus. Finally, the coefficient on average turnover is positive and significant at the five-percent level, consistent with the short sellers’ preference for liquid stocks (D’Avolio, 2002). The other variables are not statistically significant.

These findings provide new insights into the decision process of short sellers. Specifically, they appear to behave as contrarians, targeting glamour firms that have experienced a large run-up in price. They also appear to be sensitive to indicators of earnings quality and target firms with high accruals. Since, on average, firms with high momentum continue to earn positive abnormal returns (Jegadeesh and Titman, 1993), one plausible interpretation of the results is that information in financial ratios helps short sellers identify the subset of high momentum stocks whose performance is not sustainable. While prior work has documented the
A potential concern about our analysis is the extent to which our findings could be
generalized because our model is developed from a small set of potential short targets identified
by one firm. An additional concern relates to the robustness of the findings and the predictive
ability of the model beyond the specific time period of study. Ideally, analyses of short seller
behavior would be conducted using the short recommendations of all short sellers. In practice,
however, such an analysis is not possible because publicly available datasets with broad
coverage on short recommendations do not exist. However, to the extent that short sellers assume
positions on similar cues and the out of sample tests (described below) validate our model, such
concerns are mitigated.

In the next section, we describe several out of sample analyses that we conducted to test
our two hypotheses. First, we test hypothesis H1 by examining whether the firms identified by
the model as valuation shorts do indeed exhibit high short interest ratios and poor subsequent
performance. Second, to verify whether the model contains more information than the level of
short interest, we test the relative predictive ability of the model and short interest with respect to
future returns (Hypothesis H2a). We also examine the characteristics of firms identified by the
model as valuation shorts and arbitrage shorts. If the model successfully identifies arbitrage
shorts, we expect that these firms will exhibit firm characteristics that are typically associated
with arbitrage strategies (Hypothesis H2b).
B. Out of Sample Results

We now present the results of the out of sample performance of each of the four models reported in Table II. We use annual firm level data from the prediction period 1990-1996 and the coefficients of the logistic regression models obtained from the estimation period (1997-2004, presented in Table II) to calculate the predicted probability of being a valuation short target. We sort the entire population of firms each year into deciles based on the predicted probability from the model and report the results of tests using these decile portfolios.

B.1. Predicted Probability and (Observed) Short Interest Ratios

In Table III, we report the average short interest ratio for each decile portfolio over the out of sample period. The results suggest that the model does a good job of identifying firms with high short interest ratios, out-of-sample. Specifically, for model 3, the firms in decile portfolio 10, comprising firms with the highest predicted probability, have average short interest of 3.24%. In sharp contrast, the firms in decile portfolio 1, comprising firms with the lowest predicted probability, have an average short interest of 0.47%. Remarkably, the pattern of short interest shows a monotonic increase from decile 1 to decile 10. The difference in mean short interest between deciles 1 and 10 is highly statistically significant (t-statistic of 24.5). The results for models 1, 2 and 4 are similar; in every case, the monotonic pattern in short interest persists and the average short interest ratio in decile 10 is significantly higher than that in decile 1.

If our model successfully describes (at least partially) the trading behavior of short sellers, we expect that the firms identified as valuation shorts should exhibit high short interest. The results in Table III are consistent with this line of reasoning. However, while these results are suggestive, they are not conclusive as the firms in decile 10 could also be the target of arbitrage shorts. To further confirm that the model is indeed identifying valuation shorts, we
examine the pattern in future returns across the decile portfolios. If the decile 10 represents valuation shorts, we expect that the future abnormal returns of firms in decile 10 will be reliably negative. We next examine the patterns in abnormal returns for the decile portfolios.

**B.2. Abnormal Returns**

Table IV presents the subsequent stock market performance for the decile portfolios. The abnormal returns are estimated as follows. We calculate the monthly returns of an equally weighted portfolio of all firms in each decile during 1990-1996 and generate a monthly time series of portfolio returns for each decile. We match these calendar month portfolio returns with the return factors RmRf, SMB, and HML that are designed to mimic the impact of the market, firm size, and book-to-market factors on firm returns. We report the regression intercept from calendar time regressions as an estimate of the abnormal performance of each decile portfolio.

The results in Table IV, based on the three-factor model, suggest that the firms identified as valuation shorts significantly underperform. The abnormal performance of the decile 10 portfolio (model 3), comprising firms with the highest predicted probability, is –0.76 % per month (significant at the 5% level). The large magnitude of abnormal return suggests that these findings are also economically significant. In contrast, firms in the decile 1 portfolio (low predicted probability) exhibit significantly positive abnormal returns of 1.25 % per month.

Consistent with the results documented in Table III for short interest, there is a near monotonic pattern in the intercepts (abnormal returns) across the ten decile portfolios. The spread in returns between the extreme decile portfolios (decile 1 and decile 10) is 2.01 % per month. The performance is very similar for models (1), (2), and (4). The firms in the decile 10 portfolio exhibit significant underperformance for each model, ranging from –0.59 % per month to –0.74 % per month.

---

13 We downloaded the factor returns from Ken French’s website, and thank him for making the data publicly available.
% per month. The corresponding returns for decile 1 portfolio range from 1.04 % per month to 1.18 % per month.

Overall, the out of sample tests indicate that consistent with hypothesis H1, our parsimonious model for identifying valuation shorts is successful in predicting the level of short interest and future abnormal performance. In the next section, we present additional analyses to further examine our assertion (Hypothesis H2) that the short interest model can help distinguish valuation shorts from arbitrage shorts.

B.3. Additional Tests on Distinguishing Valuation Shorts from Arbitrage Shorts

To the extent that the short interest model can identify valuation shorts, the model should be able to forecast returns better than the short interest ratios, which includes both valuation and arbitrage shorts. We test this hypothesis using a two-way sort of the data. Specifically, each year, we group all firms into decile portfolios based on the short interest in October. We retain firms in the two extreme decile portfolios to maximize the spread in short interest ratio. Following the approach in Tables III and IV, we independently classify the firms into three groups each year based on the predicted probability from the short interest model. The low group (medium, high) has firms in the lowest 3 deciles (middle 4 deciles, highest 3 deciles). Thus, this two-way independent sort generates six groups (2 extreme short interest groups * 3 predicted probability groups). If the firm is primarily targeted by hedging (or arbitrage) related short sellers rather than valuation-based short sellers, we expect that the short interest ratio will be high; however, the predicted probability from the short interest model would be low, and we would not expect this group to experience negative abnormal returns. On the other hand, high predicted probability and high short interest suggests that the firm is a target of valuation shorts and these firms should experience negative returns. Finally, we expect the abnormal returns to be negative for firms
with high predicted probability but low short interest. These firms may be subject to short sale constraints that discourage short sellers from assuming positions.

Following the approach in Table IV, we estimate the intercept from calendar time regressions for an equally weighted portfolio of firms in each of the six groups. Although, we have performed the analysis for each of the four models reported in Table II, in the interests of brevity, we only report abnormal returns for model 3 in Table II (using industry-adjusted data and operating accruals). The results are qualitatively similar across all models.

From Panel A of Table V, we observe that the abnormal returns are not significantly different from zero (-0.42%) when short interest is high but the predicted probability is low. We conjecture (and provide confirmatory evidence later in panel B) that the high short interest for these firms is motivated by arbitrage related reasons. The abnormal return is -0.71% per month (significant at the one percent level) for firms with high short interest and high predicted probability, consistent with these firms being targets of valuation shorts. Interestingly, the abnormal return is –1.54% per month (significant at the one-percent level) for firms with high predicted probability but low short interest. This suggests that short sellers may not take large positions in some firms with high predicted probability, possibly due to short-sale constrains (additional evidence is provided in panel B). Finally, for firms with low short interest and low predicted probability, the abnormal return is 0.77% per month (significant at five percent level).

To test our conjecture that the high short interest for firms with low predicted probability is due to arbitrage shorts and that the low short interest for firms with high predicted probability is due to short sale constraints, we examine several characteristics for these firms in panel B of Table V. We first analyze firms with low predicted probability that have either high short interest (likely to be hedging shorts, column 2), or low short interest (column 1). We document that firms
with low predicted probability and high short interest are significantly larger and have higher turnover than firms with low predicted probability and low short interest. Importantly, 55% of firms identified as arbitrage shorts have convertibles securities outstanding and 21% of firms are members of the S&P 500 Index. In comparison, the corresponding proportion for firms with low predicted probability and low short interest are only 13% and 1%, respectively. This suggests that the high short interest for firms with low predicted probability is likely driven by arbitrage related reasons (convertibles arbitrage or index arbitrage). This finding is particularly striking because the short interest model (Table II) does not include either convertible securities outstanding or index membership as an explanatory variable. Yet, consistent with hypothesis H2, the model demonstrates the ability to differentiate between firms subject to valuation and arbitrage short selling.

We next examine firms with high predicted probability that have either low short interest (likely to be short sale constrained, column 3) or high short interest (column 4). Consistent with our conjecture that the former are short constrained, we find that these firms are significantly smaller (average size of $256 million versus $958 million) and have lower share turnover (0.28% versus 0.81%) than firms with high short interest. Consistent with the large difference in firm size, the fraction of firms that are S&P500 constituents is statistically different, although not economically large, between these two groups (0.36% and 4.26%). There is little economic difference in BM ratio or the fraction of firms with convertible securities outstanding. Hence, even though our model identifies a set of firms that subsequently experience significantly negative returns, the existence of short sale constraints likely prevents short sellers from assuming position in some of these stocks.

As a final additional test, we replicated the analysis in Panel A of Table V after excluding
firms with convertibles bonds outstanding, for the following reason. Prior research has categorized firms with high short interest and convertible bonds outstanding as arbitrage shorts and the remaining high short interest firms as valuation shorts (Asquith et al. (2005)). This raises the question - does the short interest model have incremental explanatory power over and above an indicator variable that captures the existence of convertible bonds for distinguishing between valuation and arbitrage shorts? This has added relevance because, during the time period covered by our study, only about 20% of the firms on Compustat have convertible securities outstanding. Therefore, the indicator variable approach to identifying hedging shorts would automatically classify short selling in almost 80% of the firms as valuation shorts, although some short selling in these stocks would clearly be motivated by hedging activities. Thus, we replicate the analysis for firms without convertibles bonds; a sub-sample where the convertibles indicator cannot distinguish valuation shorts from arbitrage shorts.

The results of this analysis strongly validate our model. Specifically, we document that the firms with high short interest but low predicted probability from our model experience abnormal returns of -0.05% per month (not significant). In contrast, the firms with high short interest and high predicted probability experience abnormal returns of -0.60% per month (significant at the 5% level). Thus, our model does a good job of distinguishing information and arbitrage motivated short selling, even within a selected sample of firms that excludes a large amount of arbitrage related short selling.

C. Robustness Tests

In this section, we discuss the results of several robustness tests conducted to verify

---

14 Detailed results are available from the authors upon request.
whether the results continue to hold under reasonable alterations to the empirical methodology.\textsuperscript{15} We find that our inferences remain unchanged.

First, since our methodology involves pooled time series estimation and we use all Compustat firms (excluding financials and regulated firms) as control firms, it is possible that the observations are not independent and that the significance levels are overstated. To address this issue, we replicated the analysis reported in Tables II-IV using an alternative approach to selecting control firms. Specifically, each year, we sort the universe of eligible Compustat firms on the basis of total assets and select every 10\textsuperscript{th} firm as a control firm. The results of the logistic estimation are very similar to those reported in Table II. The coefficients on accruals, BM, prior momentum, and turnover are statistically significant in each of the four models. The coefficient on SGA and SGAI are significant in two out of four models. The level of short interest for the decile of firms with the highest predicted probability is 3.41\% compared to 0.47\% for firms in the lowest decile (model 3). Finally, the out-of-sample return analysis (from model 3, reported in column 1 of Table VI) indicates that the firms in the highest decile experience abnormal returns of -0.56\% per month and firms in the lowest decile experience abnormal returns of 1.22\% per month, both significant at the 10 percent level or better. These results suggest that our results are not sensitive to the use of the universe of Compustat firms as control firms.

Second, we replicate the analysis after excluding low priced stocks (stock price below $10) from control firms since short sellers prefer to target large, liquid firms and avoid small, low priced firms. The out of sample abnormal returns for firms in the highest (lowest) decile of predicted probability, reported in column 2 of Table VI, are a statistically significant -0.58\% per month (-0.11\% per month, not significant). Thus, while the firms in the highest decile of

\textsuperscript{15} While the detailed results of the various robustness tests are not reported in tables, they are available from the authors upon request. In the interests of brevity, we only tabulate abnormal return results from estimating model 3, which uses operating accruals and industry-adjusted accounting data.
predicted probability continue to perform poorly, the superior performance of firms in the lowest decile of predicted probability (documented earlier in Table III) is confined to small and low priced firms that are expected to have large transaction costs.

Third, we conduct the out of sample analysis over a period preceding the estimation period instead of over a period following the estimation period. As discussed earlier, this is due to data availability issues, since our estimation period ends in 2004 and we do not have data subsequent to this period. However, it is possible that the research firm may have used prior period data (prior to 1998) to predict poor performers and then used this model to identify potential targets in subsequent periods. In this case, the out of sample tests may not be independent of the estimation period analysis. To mitigate concerns about such potential learning effects, we conduct the following test. We use the first three years of data (1997-1999) to estimate our model and the subsequent period data for the out of sample analysis. Since the short interest data available to us ends in December 2003, this analysis uses data ending in 2003. Even though we have limited data for model estimation and the out of sample analysis, the overall findings are consistent with those documented earlier. Specifically, in the logistic estimation, the coefficients on accruals, BM, and sales growth are significant across each of the four models. The coefficients on prior momentum, SGAII and GMI are significant in many of the models. The short interest of firms in the highest decile of predicted probability is 5.21% compared to 1.03% for firms in the lowest decile (model 3, using industry-adjusted data and operating accruals). Finally, the abnormal return of firms in the highest decile is -0.60% per month (p-value = 0.13) and for those in the lowest decile is 2.94%, significant at the one percent level (Table VI, column 3). Despite the significant drop in sample size, the monotonic pattern in short interest and abnormal returns is very similar to that documented earlier, suggesting that our findings are not
merely the outcome of learning effects. Rather, the results suggest that the short interest model performs quite well in identifying firms targeted by informed short sellers.\footnote{We also conducted an additional test where the dependent variable in the estimation period regression equals ‘1’ for one percent of firms that have the lowest returns each year, and equals ‘0’ otherwise. In out of sample tests, we find a U-shaped pattern in short interest (declining from 1.25% in decile 1 to 0.87% in decile 5, and then increasing to 2.78% in decile 10). We do not find a monotonic pattern in abnormal returns. This suggests that our methodology of identifying firms in the short database as valuation shorts yields results that are more robust in an out of sample period than an alternative methodology that naively fits on poorly performing firms.}

IV. An Application of the Short Interest Model

The short sellers’ primary interest is in identifying firms whose past performance (and hence price) is not sustainable. As a result, the model also naturally identifies firms with high predicted probability as those whose performance is expected to decline. Thus, the model can help design better empirical tests of competing theories in many settings. We present one application of the model below and suggest other avenues where our model could be applied.

Hong and Stein (1999) describe a model with two types of traders; “newswatchers”, who trade based on their private information, and “momentum traders”, who condition their trades based on past price changes. Their model predicts that, with the gradual diffusion of private information, stock returns will exhibit a short-run momentum effect (continuation) followed by longer-run reversals. Thus, stocks that are most “momentum-prone” should also be most “reversal-prone”. However, since on average, stocks with high prior momentum exhibit strong future returns over the next six-to-twelve months (Jegadeesh and Titman, 1993), how does one identify, ex-ante, among momentum stocks, those that are most prone to reversal? The evidence presented earlier suggests that the short interest model may be successful in identifying high momentum stocks whose performance is expected to reverse. We adopt the same approach that we have described earlier in section III (B.3), but sort the firms every year (1990-1996) into deciles based on the prior 12-month return, and focus on the extreme two deciles. We also
independently sort the firms into three groups – low (30%), medium (40%), and high (30%) - based on predicted probability from our model. The abnormal returns are evaluated over 12 months subsequent to portfolio formation using calendar-time regressions and the Fama and French (1993), three-factor model.

The results reported in Table VII suggest that our empirical model exhibits the ability to identify a subset of high momentum stocks, ex-ante, that are prone to reversal. In Panel A, we observe that the abnormal return of stocks in top momentum decile and high predicted probability based on our model is large and negative. Specifically, the mean abnormal return for this portfolio is -0.63% per month, with a t-statistic of -2.21. In panel B, we document that among high momentum firms, those firms identified as valuation shorts have significantly lower BM (0.29) and larger equity market value ($769 million) compared to firms that are not likely to be targeted by short sellers (BM of 0.98 and equity market value $211 million). Hence, the firms in the high momentum group are diverse and some firms (but not others) experience a price reversal. The short interest model is able to identify, ex-ante, momentum stocks that are expected to experience a price-reversal, as modeled theoretically in Hong and Stein (1999).

The short interest model has broader applications in other settings as well, wherein a similar methodology can be applied to ex-ante identify firms that are likely to experience poor performance. For example, competing theories of seasoned equity predict different market reactions to announcement of equity offerings. One set of models based on adverse selection and asymmetric information about expected future cash flows (e.g., Myers and Majluf (1984), Miller and Rock (1985)) suggest that overvalued firms will issue equity. Cooney and Kalay (1993) extend the Myers-Majluf model and argue that equity issues by some firms could elicit positive stock price reactions, if the managers’ choice-set is not restricted to only positive NPV projects.
as in the Myers-Majluf model. Since our model can help identify potentially overvalued firms that may issue equity in response to managerial private information versus other firms that may issue equity to fund profitable investments, it may allow more powerful tests of these competing hypotheses. Similarly, the model can help identify overvalued bidder firms in stock-for-stock merger transactions. Institutional and other investors could also use a similar analysis to inform their portfolio investment decisions.

V. Conclusions

We present an empirical model to distinguish information based short selling from arbitrage or hedging based short selling. The model relies on a unique database of firms identified as short targets by an independent research firm. We show that valuation indicators, such as BM ratio and prior momentum, and financial statement variables, such as accruals, gross margin, sales growth, and SG&A, can help identify potential short targets. We classify firms with high predicted probability from the model as targets of valuation-based short selling. Out of sample tests validate the model’s ability to distinguish between valuation and arbitrage shorts. Specifically, we document that firms identified as valuation shorts exhibit high short interest and poor subsequent stock performance and that firms identified as arbitrage shorts include a higher proportion with convertible bonds and a higher fraction that are in the S&P500 Index. We also document that the firms with high predicted probability but low short interest are short sale constrained, exhibiting poor liquidity and small size.

The above findings have several important implications. First, our paper provides insights into the decision process of short sellers, suggesting that the information set of short sellers is related to the information contained in valuation and financial statement indicators. Second, the
poor subsequent performance of firms in the short database and the poor out of sample performance of firms with high predicted probability provides collaborative evidence for the theoretical prediction that short sellers are informed investors. Our focus on short recommendations and on predictions from the empirical model contrasts with the prior literature that examines portfolios of firms with high short interest. Third, and perhaps most important, the short interest model can distinguish between valuation shorts and arbitrage shorts. This distinction is important because the information content of these two sources of short interest is different and the increasing use of arbitrage related short selling has contributed to a weakened relation between short interest ratio and future returns recent years. Finally, we present a simple application of the short interest model for refining the momentum strategy and identify several applications that could be addressed by future research.
References


Myers, S. C., and N. S. Majluf, 1984, Corporate financing and investment decisions when firms have information that investors do not have, *Journal of Financial Economics* 13, 187-221.


### Table I
Summary statistics

The table presents summary statistics for the sample of firms identified by an independent research firm as potential targets for short selling. The sample comprises of every report issued by the firm from September 1998, until June 2005. The table presents both unadjusted and industry-adjusted values for sample firms, where industry groups are based on two-digit SIC codes obtained from Compustat. All of the statistics are presented as of year prior to the issuance of the report. The definitions are reported in Appendix 1.

<table>
<thead>
<tr>
<th>Sample Characteristics in year -1</th>
<th>Mean</th>
<th>Median</th>
<th>Industry-adjusted Mean</th>
<th>Industry-adjusted Median</th>
<th># Obs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Performance variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equity BM ratio</td>
<td>0.3166 ***</td>
<td>0.2381 ***</td>
<td>-0.2138 ***</td>
<td>-0.2070 ***</td>
<td>53</td>
</tr>
<tr>
<td>Prior 1-yr. return</td>
<td>38.84 ***</td>
<td>9.40 **</td>
<td>40.18 ***</td>
<td>15.88 ***</td>
<td>54</td>
</tr>
<tr>
<td>ROA</td>
<td>0.0005</td>
<td>0.0350 **</td>
<td>-0.0047</td>
<td>0.0096 *</td>
<td>54</td>
</tr>
<tr>
<td><strong>Financial variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSRI</td>
<td>1.1772 ***</td>
<td>0.9524 ***</td>
<td>0.1997</td>
<td>-0.0174</td>
<td>54</td>
</tr>
<tr>
<td>GMI</td>
<td>-1.2541</td>
<td>0.9806 ***</td>
<td>-2.2570</td>
<td>-0.0381 ***</td>
<td>54</td>
</tr>
<tr>
<td>SGAI</td>
<td>1.1642 ***</td>
<td>0.9890 ***</td>
<td>0.1617</td>
<td>-0.0115</td>
<td>51</td>
</tr>
<tr>
<td>AQI</td>
<td>0.9772 ***</td>
<td>0.9956 ***</td>
<td>-0.0212</td>
<td>-0.0016</td>
<td>54</td>
</tr>
<tr>
<td>SGI</td>
<td>2.2056 ***</td>
<td>1.3358 ***</td>
<td>1.0863 ***</td>
<td>0.2725 ***</td>
<td>54</td>
</tr>
<tr>
<td>DEPI</td>
<td>1.4436 ***</td>
<td>0.9417 ***</td>
<td>0.4872</td>
<td>-0.0080</td>
<td>54</td>
</tr>
<tr>
<td>LVGI</td>
<td>0.9887 ***</td>
<td>0.9728 ***</td>
<td>-0.0114</td>
<td>-0.0183</td>
<td>54</td>
</tr>
<tr>
<td>TOTACC</td>
<td>0.2650 ***</td>
<td>0.1159 ***</td>
<td>0.2468 ***</td>
<td>0.1099 ***</td>
<td>54</td>
</tr>
<tr>
<td>OPACC</td>
<td>-0.0238</td>
<td>-0.0378 **</td>
<td>0.0383 **</td>
<td>0.0150 *</td>
<td>54</td>
</tr>
<tr>
<td><strong>Firm characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Assets</td>
<td>857.53 ***</td>
<td>405.16 ***</td>
<td>659.89 ***</td>
<td>158.24 ***</td>
<td>54</td>
</tr>
<tr>
<td>MVE</td>
<td>1089.77 ***</td>
<td>691.06 ***</td>
<td>919.84 ***</td>
<td>571.57 ***</td>
<td>54</td>
</tr>
<tr>
<td>Trading volume</td>
<td>7.29 ***</td>
<td>3.95 ***</td>
<td>6.70 ***</td>
<td>3.35 ***</td>
<td>54</td>
</tr>
<tr>
<td>Turnover (%)</td>
<td>0.69 ***</td>
<td>0.60 ***</td>
<td>0.41 ***</td>
<td>0.30 ***</td>
<td>54</td>
</tr>
</tbody>
</table>

*, ** and *** denote significance at the 10-, 5- and 1-percent level respectively.
Table II
Logistic regressions modeling the short seller’s decision to target a firm

The table presents the coefficients from logistic regressions modeling the short seller’s decision. The regressions use all firms with available data, excluding financial firms (SIC 6000-6999), utilities (4900-4999), and communications firms (4800-4899). The dependent variable is an indicator variable that takes the value of one if the firm was targeted by the short seller, and takes the value of zero otherwise. The explanatory variables include financial statement ratios identified by Beneish (1999), operating accruals, total accruals, book-to-market ratio, firm size, prior one-year return, and average daily turnover in the prior one year. All data are winsorized at the 0.5% and 99.5% levels. The specific variable definitions are in Appendix 1. Models (1) and (2) are estimated using raw data and models (3) and (4) are estimated using industry-adjusted data.

<table>
<thead>
<tr>
<th>Model</th>
<th>Unadjusted, winsorized</th>
<th>Industry-median adjusted, winsorized</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intercept</strong></td>
<td>-5.7745 ***</td>
<td>-5.9375 ***</td>
</tr>
<tr>
<td><strong>Performance variables</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BM</td>
<td>-2.3495 ***</td>
<td>-2.0908 ***</td>
</tr>
<tr>
<td>Prior 1-year return</td>
<td>0.2308 *</td>
<td>0.2289 *</td>
</tr>
<tr>
<td><strong>Financial variables</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSRI</td>
<td>-0.0200</td>
<td>0.0128</td>
</tr>
<tr>
<td>GMI</td>
<td>-0.2040</td>
<td>-0.2205 *</td>
</tr>
<tr>
<td>SGAI</td>
<td>0.7913 **</td>
<td>0.8383 **</td>
</tr>
<tr>
<td>AQI</td>
<td>-0.2267</td>
<td>-0.2929</td>
</tr>
<tr>
<td>SGI</td>
<td>0.2648 ***</td>
<td>0.5047 ***</td>
</tr>
<tr>
<td>DEPI</td>
<td>-0.2518</td>
<td>-0.2141</td>
</tr>
<tr>
<td>LVGI</td>
<td>-0.1991</td>
<td>-0.2058</td>
</tr>
<tr>
<td>TOTACC</td>
<td>2.0702 ***</td>
<td>2.3394 ***</td>
</tr>
<tr>
<td>OPACC</td>
<td>2.8047 ***</td>
<td>3.0838 ***</td>
</tr>
<tr>
<td><strong>Firm characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Turnover</td>
<td>47.6647 *</td>
<td>57.0076 **</td>
</tr>
<tr>
<td>Pseudo $R^2$ (%)</td>
<td>9.2</td>
<td>10.1</td>
</tr>
</tbody>
</table>

*, ** and *** denote significance at the 10-, 5- and 1-percent level respectively.
The table presents the mean level of short interest for decile portfolios formed on predicted likelihood of being targeted by informed short sellers. The estimation period for the model is 1997 to 2004, based on the sample of firms identified by an independent research firm as potential targets for short selling. The coefficients of the estimation model are presented in Table II. The decile portfolios are calculated based on the predicted values during the period 1990 to 1996. Specifically, for a firm with available data in every year during 1990 to 1996, the predicted likelihood of being targeted is obtained by multiplying the coefficients reported in Table II with the firm’s characteristics for the given year. Decile portfolios based on this predicted probability are constructed each year. Reported are the mean levels of short interest (as a proportion of shares outstanding) for each of the ten portfolios.

<table>
<thead>
<tr>
<th>From Table II</th>
<th>Unadjusted</th>
<th>Industry-median adjusted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model (1)</td>
<td>Model (2)</td>
</tr>
<tr>
<td>Decile 1 (Low)</td>
<td>0.510</td>
<td>0.518</td>
</tr>
<tr>
<td>2</td>
<td>0.666</td>
<td>0.605</td>
</tr>
<tr>
<td>3</td>
<td>0.792</td>
<td>0.807</td>
</tr>
<tr>
<td>4</td>
<td>0.872</td>
<td>0.898</td>
</tr>
<tr>
<td>5</td>
<td>0.966</td>
<td>0.932</td>
</tr>
<tr>
<td>6</td>
<td>1.077</td>
<td>0.989</td>
</tr>
<tr>
<td>7</td>
<td>1.134</td>
<td>1.240</td>
</tr>
<tr>
<td>8</td>
<td>1.545</td>
<td>1.376</td>
</tr>
<tr>
<td>9</td>
<td>1.947</td>
<td>1.948</td>
</tr>
<tr>
<td>Decile 10 (High)</td>
<td>3.312</td>
<td>3.466</td>
</tr>
<tr>
<td>t for High-Low</td>
<td>24.48</td>
<td>25.40</td>
</tr>
</tbody>
</table>
The table presents the abnormal returns for decile portfolios formed on predicted likelihood of being targeted by informed short sellers. The estimation period for the model is 1997 to 2004, based on the sample of firms identified by an independent research firm as potential targets for short selling. The coefficients of the estimation model are presented in Table II. The decile portfolios are calculated based on the predicted values during the period 1990 to 1996. Specifically, for a firm with available data in every year during 1990 to 1996, the predicted likelihood of being targeted is obtained by multiplying the coefficients reported in Table II with the firm’s characteristics for the given year. Decile portfolios based on this predicted probability are constructed each year. Reported are the average monthly abnormal returns for the ten portfolios, measured as the intercept from a regression of monthly portfolio returns on the three Fama-French (FF) factors RmRf, SMB, and HML.

<table>
<thead>
<tr>
<th>Decile (Low)</th>
<th>Unadjusted</th>
<th>Industry-median adjusted</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0111 ***</td>
<td>0.0104 ***</td>
</tr>
<tr>
<td>2</td>
<td>0.0074 ***</td>
<td>0.0073 ***</td>
</tr>
<tr>
<td>3</td>
<td>0.0015</td>
<td>0.0031</td>
</tr>
<tr>
<td>4</td>
<td>0.0049 **</td>
<td>0.0043 **</td>
</tr>
<tr>
<td>5</td>
<td>0.0007</td>
<td>0.0012</td>
</tr>
<tr>
<td>6</td>
<td>-0.0007</td>
<td>-0.0012</td>
</tr>
<tr>
<td>7</td>
<td>0.0004</td>
<td>-0.0002</td>
</tr>
<tr>
<td>8</td>
<td>-0.0002</td>
<td>-0.0024</td>
</tr>
<tr>
<td>9</td>
<td>-0.0042 *</td>
<td>-0.0044 **</td>
</tr>
<tr>
<td>10 (High)</td>
<td>-0.0074 **</td>
<td>-0.0059 *</td>
</tr>
</tbody>
</table>

* *, ** and *** denote significance at the 10-, 5- and 1-percent level respectively.
Table V
Abnormal return and firm characteristics, categorized by the predicted probability and the aggregate level of short interest

The table presents the abnormal returns (panel A) and firm characteristics (panel B) for subsets of data categorized using a two-way sort based on the predicted likelihood of being targeted by informed short sellers and the actual level of short interest. For the period from 1990 to 1996, the firms are sorted into deciles every year based on the level of short interest (normalized by the number of shares outstanding). We retain the two extreme deciles to maximize the spread in short interest. We also independently sort the firms into three groups based on the predicted likelihood of being targeted by informed short sellers (deciles 1-3, 4-7, and 8-10). The predicted likelihood is obtained by multiplying the coefficients (from the estimation period logistic regressions) with the firm’s characteristics for the given year. Decile portfolios based on this predicted probability are constructed each year. In panel A, we report the average monthly abnormal returns for the six portfolios, measured as the intercept from a regression of month portfolio returns on the three Fama-French (FF) factors RmRf, SMB, and HML. In panel B, we report the average firm characteristics for these groups. The significance levels in panel B test for differences between the low (decile 1) and high (decile 10) actual short groups, holding the predicted probability constant.

Panel A: Abnormal return

<table>
<thead>
<tr>
<th>Predicted probability</th>
<th>Low (deciles 1 - 3)</th>
<th>Medium (deciles 4 - 7)</th>
<th>High (deciles 8 - 10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low short interest (decile 1)</td>
<td>0.0077 **</td>
<td>0.0013</td>
<td>-0.0154 ***</td>
</tr>
<tr>
<td>High short interest (decile 10)</td>
<td>-0.0042</td>
<td>-0.0069 ***</td>
<td>-0.0071 ***</td>
</tr>
</tbody>
</table>

Panel B: Firm characteristics

<table>
<thead>
<tr>
<th>Predicted probability</th>
<th>Low (deciles 1-3)</th>
<th>High (deciles 8-10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short interest</td>
<td>Low (decile 1)</td>
<td>High (decile 10)</td>
</tr>
<tr>
<td>MVE ($ mill)</td>
<td>70.59</td>
<td>816.90 ***</td>
</tr>
<tr>
<td>BM</td>
<td>0.93</td>
<td>0.82 ***</td>
</tr>
<tr>
<td>Average turnover (%)</td>
<td>0.13</td>
<td>0.46 ***</td>
</tr>
<tr>
<td>% with cvt. sec.</td>
<td>12.75</td>
<td>54.54 ***</td>
</tr>
<tr>
<td>% in S&amp;P 500 Index</td>
<td>0.95</td>
<td>20.78 ***</td>
</tr>
</tbody>
</table>

*, ** and *** denote significance at the 10-, 5- and 1-percent level respectively.
#, ## and ### denote significance at the 10-, 5- and 1-percent level respectively of the difference between high and low short interest groups.
Table VI  
Sensitivity analysis of post-event abnormal returns 
categorized by the predicted probability from the short interest model

The table presents the abnormal returns for decile portfolios formed on predicted likelihood of being targeted by informed short sellers. We estimate three alternative specifications of the logistic regression model, similar to those in Table II – including only every tenth firm (ranked yearly by total assets) as control firms, excluding firms if the share price is below $10, and using 1997-1999 as the estimation period and 2000-2003 as the out of sample period. The decile portfolios are calculated based on the predicted values during the out of sample period. Specifically, for a firm with available data in every year in the out of sample period, the predicted likelihood of being targeted is obtained by multiplying the coefficients (from the estimation period logistic regressions) with the firm’s characteristics for the given year. Decile portfolios based on this predicted probability are constructed each year. Reported are the average monthly abnormal returns for the ten portfolios, measured as the intercept from a regression of month portfolio returns on the three Fama-French (FF) factors RmRf, SMB, and HML.

<table>
<thead>
<tr>
<th>Abnormal returns</th>
<th>Include every 10th firm as control</th>
<th>Exclude if share price &lt; $10</th>
<th>Estimation period 1997-1999</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0122 ***</td>
<td>-0.0011</td>
<td>0.0294 ***</td>
</tr>
<tr>
<td>2</td>
<td>0.0076 ***</td>
<td>0.0010</td>
<td>0.0161 **</td>
</tr>
<tr>
<td>3</td>
<td>0.0044 *</td>
<td>-0.0002</td>
<td>0.0161 *</td>
</tr>
<tr>
<td>4</td>
<td>0.0036</td>
<td>0.0007</td>
<td>0.0117 **</td>
</tr>
<tr>
<td>5</td>
<td>0.0015</td>
<td>0.0009</td>
<td>0.0075 *</td>
</tr>
<tr>
<td>6</td>
<td>0.0008</td>
<td>-0.0019</td>
<td>0.0023</td>
</tr>
<tr>
<td>7</td>
<td>-0.0006</td>
<td>0.0009</td>
<td>0.0026</td>
</tr>
<tr>
<td>8</td>
<td>-0.0026</td>
<td>-0.0023</td>
<td>0.0017</td>
</tr>
<tr>
<td>9</td>
<td>-0.0045**</td>
<td>-0.0039 **</td>
<td>0.0008</td>
</tr>
<tr>
<td>Decile 10 (High)</td>
<td>-0.0056 *</td>
<td>-0.0058 ***</td>
<td>-0.0060</td>
</tr>
</tbody>
</table>

*, ** and *** denote significance at the 10-, 5- and 1-percent level respectively.
Table VII  
Abnormal return and firm characteristics, categorized by the predicted probability and prior momentum

The table presents the abnormal returns (panel A) and firm characteristics (panel B) for subsets of data categorized using a two-way sort based on the predicted likelihood of being targeted by informed short sellers and prior momentum. For the period from 1990 to 1996, the firms are sorted into deciles every year based on the 12-month momentum. We retain the two extreme deciles to maximize the spread in momentum. We also independently sort the firms into three groups based on the predicted likelihood of being targeted by informed short sellers (deciles 1-3, 4-7, and 8-10). The predicted likelihood is obtained by multiplying the coefficients (from the estimation period logistic regressions) with the firm’s characteristics for the given year. Decile portfolios based on this predicted probability are constructed each year. In panel A, we report the average monthly abnormal returns for the six portfolios, measured as the intercept from a regression of month portfolio returns on the three Fama-French (FF) factors RmRf, SMB, and HML. In panel B, we report the average firm characteristics for these groups. The significance levels in panel B test for differences between the low (deciles 1-3) and high (deciles 8-10) predicted probability groups, holding the prior momentum constant.

<table>
<thead>
<tr>
<th>Panel A: Abnormal return</th>
<th>Predicted probability</th>
<th>Low (deciles 1 - 3)</th>
<th>Medium (deciles 4 - 7)</th>
<th>High (deciles 8 - 10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low momentum (decile 1)</td>
<td>0.0102</td>
<td>0.0028</td>
<td>-0.0080</td>
<td></td>
</tr>
<tr>
<td>High momentum (decile 10)</td>
<td>0.0061</td>
<td>0.0038</td>
<td>-0.0063 **</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Firm characteristics</th>
<th>Predicted probability</th>
<th>Prior momentum</th>
<th>Low (decile 1)</th>
<th>High (decile 10)</th>
<th>Low (dec. 1-3)</th>
<th>High (dec. 8-10)</th>
<th>Low (dec. 1-3)</th>
<th>High (dec. 8-10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MVE ($ mill)</td>
<td></td>
<td>Low (dec. 1)</td>
<td>95.1</td>
<td>179.0</td>
<td>210.6</td>
<td>769.4***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BM</td>
<td></td>
<td>High (dec. 8-10)</td>
<td>0.73</td>
<td>0.26***</td>
<td>0.98</td>
<td>0.29***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average turnover (%)</td>
<td></td>
<td>Low (dec. 1-3)</td>
<td>0.27</td>
<td>0.55***</td>
<td>0.31</td>
<td>0.60***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% with cvt. sec.</td>
<td></td>
<td>High (dec. 8-10)</td>
<td>26.0</td>
<td>19.3***</td>
<td>18.8</td>
<td>19.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% in S&amp;P 500 Index</td>
<td></td>
<td>Low (dec. 1-3)</td>
<td>0.86</td>
<td>0.76</td>
<td>2.69</td>
<td>2.55</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*, ** and *** denote significance at the 10-, 5- and 1-percent level respectively.  
#, ## and ### denote significance at the 10-, 5- and 1-percent level respectively of the difference between high and low predicted probability groups.
Appendix 1  
Variable definitions

The table presents the definitions of the variables used in the analysis. All #s pertain to the data item numbers from the Compustat annual files.

**Performance variables**

Equity BM ratio  
#60_t / [#25_t * #199_t]

Prior 1-yr. return  
Raw return from October of year ‘t’ through September of year ‘t+1’

Return on assets  
#18_t / #6_t

**Financial variables**

DSRI (Days in sales rec. index)  
[#2_t / #12_t] / [#2_{t-1} / #12_{t-1}]

GMI (Gross margin index)  
[(#12_{t-1} - #41_t) / #12_t] / [(#12_t - #41_{t-1}) / #12_{t-1}]

SGAI (Selling, general, and admin. expenses index)  
[#189_t / #12_t] / [#189_{t-1} / #12_{t-1}]

AQI (Asset quality index)  
[(1-(#4_t + #8_{t-1})) / #6_t] / [(1-(#4_{t-1} + #8_{t-1})) / #6_{t-1}]

SGI (Sales growth index)  
#12_t / #12_{t-1}

DEPI (Depreciation index)  
[(#14_{t-1} - #65_t) / (#14_t - #65_{t-1} + #8_t)] / [(#14_{t-1} - #65_{t-1}) / (#14_t - #65_{t-1} + #8_{t-1})]

LVGI (Leverage index)  
[#5_t + #9_t] / #6_t / [(#5_{t-1} + #9_{t-1}) / #6_{t-1}]

TOTACC (Total accruals)  
[#18_t - #308_t - #311_t] / [(#6_t + #6_{t-1}) / 2]

OPACC (Operating accruals)  
[#18_t - #308_t] / [ (#6_t + #6_{t-1}) / 2]

**Firm characteristics**

Total Assets  
#6_t

Equity market value  
Share Price * Number of shares outstanding at September month end, from CRSP.

Trading volume  
The average of the daily trading volume (price * number of shares traded) from October of year ‘t’ to September of year ‘t+1’.

Turnover  
The average of the daily turnover (number of shares traded / Number of shares outstanding) from October of year ‘t-1’ to September of year ‘t’.