The Role of Fundamental Analysis in Information Arbitrage: Evidence from Short Seller Recommendations

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: April 2007

We thank David Aboody, Hank Bessembinder, Alex Butler, Jennifer Francis, Ryan LaFond, Per Olsson, Kristian Rydqvist, Ajai Singh, Mary Stanford, Rex Thompson, Brett Trueman, Mohan Venkatachalam 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. An earlier version of the paper was titled "On Distinguishing between Information and Arbitrage Motivated Short selling".

The Role of Fundamental Analysis in Information Arbitrage: Evidence from Short Seller Recommendations

Abstract

We examine whether information arbitrageurs attempt to exploit the return predictability in valuation and fundamental signals. Using a unique database of short sale recommendations, we document that firm fundamentals, such as accruals, sales growth, gross-margin and SG&A, and valuation indicators, such as book-to-market ratio and return momentum, contain valuable information correlated with the trading behavior of short sellers. We show that our empirical model explaining short seller recommendations is successful in predicting both short interest and future returns for a broader sample in an *out-of-sample* period. We present an important application of the model in distinguishing between valuation and arbitrage-motivated short selling. Overall, these findings present additional insights into the decision process of short sellers and validate the importance of fundamental analysis in the information arbitrage process.

Keywords: Information arbitrage; short selling, fundamental analysis, stock returns.

A significant body of research has shown that various valuation and financial statement indicators predict subsequent returns. For example, Lakonishok, Shleifer and Vishny (1994) show that valuation multiples, such as P/E ratio and B/M ratio, predict future returns. Sloan (1996) documents a strong association between accruals and future returns. Abarbanell and Bushee (1997, 1998) document that information contained in fundamental analysis is related to both future earnings and returns.[1] These studies argue that investors can earn "abnormal" returns by trading on various signals of financial performance, as the market fails to fully incorporate the information in historical financial data into prices in a timely manner. Although the evidence on return predictability is strong, there is little *direct* evidence on whether information arbitrageurs attempt to exploit such predictability. In this study, we examine a unique database of short seller recommendations to better understand the trading strategies of information arbitrageurs.

A *direct* examination of the trading behavior of information arbitrageurs is important because there is significant debate in the literature on the interpretation of return predictability. One set of studies suggests that predictability patterns tend to be exaggerated due to data errors and biased methodologies (Kothari (2001), Leone et al. (2005)). Another set of papers has attempted to reconcile the predictable return behavior within a risk based framework (Fama and French (1995), Francis et al. (2005)). A third set of papers argues that return predictability is difficult to exploit as it is confined to stocks where arbitrage is costly and risky (Ali et al. (2003), Mashruwala et al. (2004)). To the extent that return predictability can be fully explained by methodological biases, risk or market frictions, we would expect to see no clear relation between the trading behavior of information arbitrageurs and the information contained in fundamental

[1] Among related work, Lev and Thiagarajan (1993) show that financial statement ratios are related to contemporaneous returns. Abarbanell and Bushee (1998) document that a trading strategy based on a combination of several financial statement variables generates significant abnormal returns. Piotroski (2000) shows that fundamental analysis can help refine the value investment strategy. Beneish, Lee and Tarpley (2001) show that fundamental variables are useful in separating losers from winners among the subset of extreme performers.

signals. On the other hand, a finding that fundamental signals are related to the trading behavior would support the idea that information arbitrageurs assume positions in an attempt to exploit the return predictability in fundamental signals.

We focus on short sellers as representing an important group of information arbitrageurs for two reasons. First, earlier studies examining trading strategies based on a valuation or fundamental signals have documented that a large component of abnormal returns is earned on the short position in the hedge portfolio. For example, results in Beneish and Vargus (2002) suggest that the returns to the accruals anomaly are due to overpricing of income-increasing accruals. Second, and more importantly, extant theoretical and empirical literature supports the notion that short sellers play an important role in the information arbitrage process. Theoretical models predict that short sellers are more likely to be informed traders, as short selling costs tend to disproportionately discourage liquidity traders from selling short (Diamond and Verrecchia (1987)). In support of theoretical predictions, prior empirical research documents that heavily shorted firms subsequently exhibit negative abnormal performance.[2] Yet, despite their characterization as informed arbitrageurs, we know relatively little about their trading behavior, mainly due to lack of publicly available data on short sellers' recommendations or positions.

A few studies in the literature have examined the characteristics of firms with high reported short interest; the motivation being that reported short interest proxies for short selling demand from informed traders targeting overvalued securities. These studies find that short interest is related to valuation multiples, such as book-to-market ratio (Dechow et al (2001)), accruals (Desai et al (2006), Hirshleifer et al (2005)), post-earnings announcement drift (Cao et

[2] Asquith and Meulbroek (1995), Desai, Ramesh, Thiagarajan and Balachandran (2002), Christophe, Ferri and Angel (2004), Boehmer, Jones and Zhang (2005), and Diether, Lee and Werner (2005), among others, document a negative association between short interest and future returns. Dechow, Hutton, Meulbroek, and Sloan (2001) discuss the various legal and institutional constraints related to short selling.

al. (2006)) and firm liquidity (D'Avolio (2002)). The demand for short selling however could be motivated either by a pessimistic opinion on firm valuation (valuation shorts) or by various arbitrage or hedging strategies (arbitrage shorts).[3] Asquith, Pathak and Ritter (2005) argue that the reported short interest in the U.S., which aggregates both valuation and arbitrage shorts, is an imprecise proxy for short selling demand by information arbitrageurs, especially in recent years, due to the increasing use of various "market neutral" investment strategies by institutions.

Our study is distinguished from prior work on short interest partly because we examine the information content of a much broader set of fundamental signals for short sellers, but mainly because we provide *direct* evidence on the trading behavior of information arbitrageurs by examining a unique database of short sale recommendations. The *short database* includes *every* recommendation issued by an independent research firm for its institutional clients since its inception in 1998 until June 2005. Importantly, these recommendations are only motivated by a perceived overvaluation of the firm's stock and are unrelated to arbitrage strategies, thus providing a more precise signal of valuation-motivated shorting demand. Consistent with this reasoning, we find that firms in the short database experience mean raw return (market-adjusted return) of -4.03% (-4.89%) in the month when the short report was issued and –9.71% (-15.02%) in the subsequent 12 months. We document that the sample firms report a disproportionately large number of negative news events during the subsequent 12 months – the most common being reporting lower than expected earnings, lowering guidance on future earnings / sales, and analyst downgrades. These findings suggest that the research firm is successful in identifying

[3] 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 (or ETFs) and the component stocks. Mitchell, Pulvino, and Stafford (2004) conclude that, for firms with pending mergers or convertible bonds outstanding, the short sales due to arbitrage activities is likely to be much larger than valuation motivated shorts. Boehmer, Jones, and Zhang (2006) examine a proprietary system order dataset for NYSE stocks and report that over 25% of the total short sale daily volume in a stock can be attributed to program trading.

short targets whose financial performance is expected to deteriorate significantly during the following year.

Based on the short database, we build a parsimonious model of the decision process of short sellers based on context-specific financial performance measures. The short sellers' primary motive lies in identifying firms that are expected to perform poorly. Prior research on short seller behavior has focused primarily on either valuation ratios or on signals generated by accruals. We expand the set of explanatory variables by integrating evidence from both the accounting and finance literatures and examine the incremental explanatory power of each variable on short seller behavior. We include financial statement variables that are identified by Beneish (1999) as being related to earnings manipulation. We also examine the importance of valuation multiples and firm characteristics that have been identified by prior research.

The results of logistic regression analysis, estimated over the sample period 1997-2004, indicate that accounting information, especially indicators of earnings quality, plays a critical role in identifying promising short targets. Short sellers are more likely to target firms with high accruals that have also experienced large increases in sales, SG&A expenses, and gross profit margin. These findings are intuitive, as a significant growth in sales and gross margin, coupled with high accruals, may signal that the earnings growth is not sustainable. Our results also suggest that short sellers target firms with low book-to-market ratio and high one-year return momentum, suggesting that short sellers tend to be contrarians, betting that the strong past performance of targeted firms will revert quickly. Finally, consistent with short sellers' preference for liquid securities to minimize short squeeze risk, we find that short recommendations tend to be in relatively liquid stocks. Overall, these findings suggest that the return predictability associated with fundamental signals represents an economically meaningful opportunity that informed arbitrageurs attempt to exploit. Thus, our examination of short seller recommendations documents a direct link between the literature on the predictive ability of fundamental signals and the literature on the trading strategies of information arbitrageurs.

We validate the short interest model by examining both short interest and future returns in an *out-of-sample* period (1990-1996). Specifically, we sort firms into decile portfolios constructed each year based on the predicted probability from the estimated short interest model. We observe a monotonic trend in short interest ratio across decile portfolios, increasing from about 0.5% for firms in the lowest decile to over 3.2% for firms in the highest decile. This finding is particularly noteworthy, as the model does not include short interest as an explanatory variable. Yet, the model is able to predict short interest in an *out-of-sample* period. A monotonic pattern is also observed for average monthly abnormal returns. The intercept from a regression of monthly portfolio returns on the Fama-French factors decreases from 1.3% for firms in the lowest decile to about –0.80% 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 a single research firm, the model is successful in forecasting both short interest and future returns for a broader sample in an *out-of-sample* period.[4] Importantly, these findings validate the importance of accounting-based signals for a broad group of information arbitrageurs and suggests that the selection criteria is correlated across short sellers and, to some extent, constant across time.

The short interest model can help design better empirical tests in many different settings. As an application, we present an approach to distinguish between valuation shorts, which are

[4] As far as we are aware, no publicly available data set exists with broad coverage on short recommendations or positions of short sellers. Because the research firm examined here faces time and resource constraints, we expect that many firms that are suitable targets for short selling are not identified as such in our database, thus reducing the predictive power of the short interest model. Despite this limitation, we find that the model can predict both short interest and future returns for a broad sample in an out-of-sample period.

bearish bets undertaken by information arbitrageurs, and arbitrage shorts, which are motivated by various hedging strategies. The short interest data reported in the United States aggregate both valuation shorts and arbitrage shorts. Although the information content with respect to future returns of the two sources of short selling is quite different, prior research on valuation shorts had to rely on aggregate short interest as an empirical proxy, 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.[5] Confirming this trend, Asquith, Pathak and Ritter (2005) document that the percentage of shares held short has increased significantly over 1988 to 2002. More importantly, they document a weakening trend in the relation between short interest and future returns, suggesting that the aggregate short interest is an imprecise proxy for valuation shorts, especially in recent years. Given the continuing growth in institutions that engage in arbitrage strategies, this problem is expected to become even more acute over time, suggesting that distinguishing between valuation shorts and arbitrage shorts presents an important avenue for current research.

We show that our model has the ability to differentiate valuation shorts from arbitrage shorts, ex ante. Specifically, for firms in the highest and lowest decile portfolios based on (observed) short interest ratios, we categorize stocks into three groups based on the predicted probability from the model (30%, 40%, and 30%), 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

[5] 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 $39 billion to $972 billion (Agarwal and Naik (2005)).

low predicted probability as arbitrage 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 (glamour firms), are less likely to have convertible bonds outstanding, and are less likely to be included in the S&P 500 index. The latter finding is quite striking, as the short interest model does not include either convertible bonds or index membership as explanatory variables. Consistent with theoretical predictions from the short interest literature, the incremental forecasting power of short interest with respect to future returns increases when valuation shorts are distinguished from arbitrage shorts. Thus, we illustrate the usefulness of conducting fundamental analysis in context.

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 distinguishing information shorts from arbitrage shorts 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 on short interest and develop the hypotheses that underlie the empirical tests that follow.

A. Performance and Firm Characteristics Associated with Short Selling

Prior empirical work broadly supports Miller's (1977) prediction that stock prices could be biased upward when short constraints exist and investor's beliefs about security value are widely dispersed (see, for example, Boehme, Danielson and Sorescu (2005), Chen, Hong and Stein (2002), Jones and Lamont (2002), Nagel (2005), among others). Extant empirical evidence also supports Diamond and Verrecchia (1987))'s prediction that short sellers are informed traders (see, for example, Senchack and Starks (1993) and Arnold, Butler, Crack, and Zhang (2005)). Several papers (see footnote 2) document that heavily shorted firms subsequently experience poor returns. Along similar lines, Francis, Venkatachalam, and Zhang (2005) find that downward revisions in analysts' forecasts are more severe for firms with high unexpected short interest.

Researchers have also examined the characteristics of firms with high reported short interest. Dechow, Hutton, Meulbroek and Sloan (2001) document that firms with high short interest trade at high multiples of price-to-fundamentals ratios. Desai, Krishnamurthy and Venkataraman (2006) find that short sellers increase positions in firms that subsequently restate earnings and, in particular, target restating firms with high accruals, suggesting that short sellers are sensitive to earnings quality. Hishleifer, Teoh and Yu (2005) and Cao, Dhaliwal and Kolasinski (2006) provide evidence suggesting that short sellers assume positions based on accounting anomalies, including the post earnings announcement drift. D'Avolio (2002) examines a unique database of stock loans and reports that short sellers have difficulty in borrowing smaller and less liquid securities.

B. Valuation versus Arbitrage Motivated Short Sales

In an attempt to discriminate between valuation and arbitrage shorts, Asquith et al. (2005) classify high short interest firms with convertible bonds outstanding as being arbitrage motivated and the remaining high short interest firms as being valuation motivated. 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, short interest in many firms with no convertible bonds would be arbitrage motivated, and vice-versa. Asquith et al. (2005) acknowledge that their classification is imprecise, and encourage researchers to develop better approaches to addressing the problem; a task that we undertake in this study.

Boehmer, Jones and Zhang (2006) examine a proprietary system order dataset for NYSE stocks that allows them to distinguish between institutional and proprietary short sales that were part of a program trade. They document that short selling that were not part of a program trade are informative about future returns, while short selling associated with program trade are not, emphasizing once again the importance of disentangling the components of short interest.

C. Hypotheses

Our objective is to better understand the trading strategies of short sellers, and in particular, examine whether short sellers attempt to exploit the return predictability of valuation and financial variables. Thus, we hypothesize that short sellers use fundamental (accounting based) information to identify target firms. If the short interest model, based on short recommendations from a single research firm, sufficiently describes the selection criteria of the broader group of short sellers, then the model should forecast both short interest and future returns in an *out-of-sample* period. Finally, we examine if the model can distinguish between valuation and arbitrage shorts. Thus, firms identified as valuation shorts should experience poor subsequent returns. In contrast, firms identified as arbitrage shorts should not exhibit poor returns but instead should exhibit characteristics associated with arbitrage strategies, such as index membership and convertible bonds outstanding. The following hypotheses summarize these arguments:

H1: Short sellers use fundamental (accounting based) variables to identify target firms.

H2: The short interest model should forecast short interest and abnormal returns in an out-of-sample period.

H3: 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 due to fundamental reasons and are not related to arbitrage strategies. 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 initial sample in the short database consists of 67 firms identified in the reports during 1998-2005 as promising short candidates. The targeted firms have experienced significant price increases during the year prior to the issuance of the short report. The mean buy and hold return is 77.9%, and the mean market-adjusted return is 72.4% (Table I, Panel A). 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 average market-adjusted return is -4.89% (t-statistic of -2.50). 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 15.0% (t-statistic of -2.18). Even the mean and median raw returns of the sample firms are negative over the subsequent 12 months (-9.71% and -18.32%, respectively).

To better understand the reasons for poor stock market performance, we searched for significant negative news events for sample firms in *Factiva* during the 12 months period after the issuance of the report. Panel B of Table I summarizes the findings. Of the 67 sample firms, we could identify some form of 'bad news' events for 42 firms. The most frequent negative news related to 'reported lower than expected earnings' (24 events), 'analyst downgrades' (23 events) and 'lowered guidance on future earnings / sales' (20 events). Eight sample firms came under some form of regulatory scrutiny and seven firms were identified to be associated with accounting concerns. Not surprisingly, some firms had multiple types of bad news events. For example, Rite Aid Corp, identified in a short report on November13 1998, reported (1) on March 12, 1999, that earnings will miss analysts' estimates, (2) on March 13, 1999 that its accounting practices are under review by federal regulators, and (3) on June 1, 1999, that it is restating the last three years of earnings. During the one-year period after the issuance of the report, Rite Aid Corp's stock price dropped by 83%.

At the same time, we find no report of negative news event for 25 of the 67 firms in the short database. Indeed, some firms in the short database performed spectacularly well during the subsequent 12 months. For example, Extended Stay America, identified in a short report on December 18, 2003, was acquired by the Blackstone Group at a significant (24%) premium, in March 2004, resulting in price run-up of 54% subsequent to report issuance. Thus, it is clear that the research firm is not always successful in identifying poor performers. However, on average, the firms in the short database experience a significant reversal in their fortunes subsequent to issuance of the report and exhibit a disproportionate share of negative news events, mainly related to fundamental variables. These findings suggest that the short database represents an appropriate sample for modeling the behavior of information arbitrageurs.

In the empirical analysis, 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.[6] 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 firms if CRSP data is available, thus eliminating Compustat observations that relate to non-public firms and subsidiaries. The CRSP screen reduces the

[6] 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.

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, 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 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.[7] Further, for sample firms identified as short targets during the period from October 1999 to September 2000, the pre-event

[7] 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.

financial data is from the 1998 fiscal year.[8]

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:

$$S_i = L\,(\alpha + \beta_i * \mathbf{X}_i + \epsilon_i) \qquad (1)$$

where S_i is a indicator variable that equals '1' for sample firms identified in the short database and equals '0' for control firms, X_i is the vector of variables selected to proxy for the information set of short sellers, ϵ_i is the residual error term, and L indicates that the model is based on a logistic regression.

The selection of the explanatory variables, motivated by the literature on contextual fundamental analysis (see Beneish et al (2001)), relies on extant short interest literature (discussed in Section I). The short seller is interested in identifying firms whose performance is likely to reverse. We include financial statement variables that are designed to capture the quality of earnings. We include two measures of accruals, namely total accruals (*TOTACC*) and

[8] 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.

operating accruals (*OPACC*).[9] 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 revenue inflation or relaxation of credit policy to generate higher sales, we expect that DSRI will be positively related to being identified in the short report. We include the gross margin index since improved margins accompanied by high accruals might suggest that the increased margins are not 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 identified in the short report.

Since financial statement ratios are likely to vary across industries, we estimate the

[9] 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)).

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. We also include valuation multiples and firm characteristics identified in the prior literature as explanatory variables. The valuation indicators are 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.[10] 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 $_{t}$.[11]

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.

[10] 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.

[11] The detailed definition of all the variables is presented in Appendix 1.

Table II 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%. 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 II, 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.

III. Results

In this section, we present the main results of the study. We first describe our estimation period results, where the model is estimated using a broad set of valuation and financial statement indicators over the Compustat years 1997-2004 using a logistic regression (Hypothesis H1). To test for 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) short interest and the abnormal stock performance across decile portfolios (Hypothesis H2). 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 (2005). In the application, we test whether our model can distinguish between valuation shorts and arbitrage shorts (Hypothesis H3).

A. Estimation Period Results

In Table III, we report the coefficients of logistic regressions that relate short recommendations to various fundamental and valuation signals. 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.[12] 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 III.

[12] To minimize the effect of outliers, all variables are winsorized at the 0.5% and 99.5% level.

Focussing first on financial statement variables, the results suggest that short recommendations are sensitive to the 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 research firm targets firms with an increase in gross margin, ceteris paribus. These findings suggest that the return predictability associated with fundamental signals represents an economically meaningful opportunity that informed arbitrageurs attempt to exploit.

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 believes in a contrarian investment strategy, targeting glamour firms that have experienced a large 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 identifies stocks whose price has been bid up due to such extrapolation and might be expected to reverse. 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. Importantly, they appear to be particularly sensitive to information contained in accounting ratios, and in particular, target firms with poor earnings quality. Since, on average, firms with high momentum continue to earn positive abnormal returns (Jegadeesh and Titman, 1993), one plausible interpretation of our findings is that the information in financial ratios helps short sellers in identifying the subset of high momentum stocks whose performance is not sustainable, illustrating the importance of context-based fundamental analysis. Thus, these findings provide a direct empirical link between the predictive ability of fundamental signals and the trading strategies of information arbitrageurs.

B. Out of Sample Results

One potential concern is the extent to which our findings could be generalized, as our model is developed from a small set of short recommendations from a single research firm. Specifically, the concern relates to whether the model can predict the trading behavior of a broader group of valuation-motivated short sellers 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 or positions 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.

To validate the model, we test whether the firms identified by the model as valuation shorts do indeed exhibit high short interest ratios and poor subsequent performance in an out-of – sample period. If the model sufficiently describes the trading behavior of a broader group of short sellers, then the firms identified by the model as valuation short targets should exhibit high short interest levels. Further, to the extent that the high short interest reflects valuation shorts and not arbitrage shorts, the future returns of firms identified by the model as valuation short targets should be reliably negative (hypothesis H2).

Table IV and Table V present the results of the out-of-sample performance of each of the four short interest models reported in Table III. 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 III) to calculate the predicted probability of being a valuation-motivated 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.

In Table IV, 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. These findings are particularly noteworthy, as the model does not include short interest as an explanatory variable in the logistical regression.

To further confirm that the model is indeed identifying valuation shorts, we report on the subsequent stock market performance for the decile portfolios in Table V, 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 V, based on the three-factor model, suggest that the firms identified as valuation shorts significantly under perform.[13] 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 IV 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. The corresponding returns for decile 1 portfolio range from 1.04 % per month to 1.18 % per month.

Several important implications emerge from the finding that the short interest model can predict both short interest and future returns in an out-sample period (hypothesis H2). First, the

[13] We downloaded the factor returns from Ken French's website, and thank him for making the data publicly available.

results suggest that the model captures the trading behavior of a broader group of valuation-based short sellers. Second, short sellers in aggregate attempt to exploit the return predictability in valuation and fundamental signals. Third, short sellers in aggregate appear to trade on a similar set of fundamental signals, which to some extent is constant across time.

C. Robustness Tests

In this section, we discuss several robustness tests that were conducted to verify that the results continue to hold under reasonable alterations to the empirical methodology.[14] 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 III-V 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 10th firm as a control firm. The results of the logistic estimation are very similar to those reported in Table III. The coefficients on accruals, BM, prior momentum, and turnover are statistically significant in each of the four models. The coefficients 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

---

[14] 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.

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 predicted probability continue to perform poorly, the superior performance of firms in the lowest decile of predicted probability (documented earlier in Table V) is confined to small and low priced firms that are expected to have large transaction costs.

Third, we elected to 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 have limited 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, SGAI 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 valuation-motivated short sellers.[15]

IV. Application: Distinguishing Valuation Shorts from Arbitrage Shorts

As described earlier, the short interest data in the United States aggregates both valuation and arbitrage shorts. Asquith et al (2005) suggest that the increased incidence of arbitrage shorts in recent years has caused the relation between short interest and future returns to weaken. In this section, we present an application of the short interest model to distinguish between valuation and arbitrage shorts (hypothesis H3), 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 IV and V, 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

[15] 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 valuation shorts using short recommendations yields results that are more robust in an out of sample period than an alternative methodology that naively fits on poorly performing firms.

two-way independent sort generates six groups (2 extreme short interest groups * 3 predicted probability groups).

The evidence thus far suggests that the short interest model can successfully identify firms targeted for valuation reasons. Thus, if the reported short interest is high and the predicted probability from the model is also high, we classify the short interest in those firms as being valuation shorts. On the other hand, we classify firms with high short interest but low predicted probability from the model as arbitrage shorts. To test our classification, we examine both future returns and firm characteristics of the six groups. Firms classified as valuation shorts should experience significant negative future returns. Further, firms classified as arbitrage shorts should not exhibit negative future returns but should exhibit firm characteristics associated with arbitrage trading, such as index membership or convertibles bonds outstanding. We also examine characteristics of firms with low short interest but high predicted probability. We hypothesize that these firms are subject to short sale constraints that discourage short sellers from assuming short positions in these firms.

Following the approach in Table V, 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 III, in the interests of brevity, we only report abnormal returns for model 3 in Table III (using industry-adjusted data and operating accruals). The results are qualitatively similar across all models.

From Panel A of Table VII, we note that abnormal returns are not significantly different from zero (0.08%) for firms with high short interest but low predicted probability. We conjecture (and provide confirmatory evidence in panel B) that the high short interest for these firms is likely driven by arbitrage strategies. On the other hand, the abnormal return is -0.71% per month (significant at one percent level) for firms with high short interest and high predicted probability, consistent with short interest in these firms being valuation shorts. The differences between the two groups are not the outcome of a finer sort on the level of short interest, as the average short interest in both groups is similar (7.43% and 8.74%, respectively). Interestingly, firms with high predicted probability but low short interest exhibit an abnormal return of –1.08% per month. We conjecture that the short interest is low for these firms due to short-sale constraints. Finally, for firms with low short interest and low predicted probability, the abnormal return is 1.19% per month (significant at five percent level).

We examine several firm characteristics in panel B of Table VII. We first examine firms with low predicted probability that have either high short interest (arbitrage shorts, column 2), or low short interest (column 1). 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, 56% of the firms with low predicted probability and high short interest have convertibles securities outstanding and 20% are members of the S&P 500 Index. In comparison, the corresponding proportions for firms with low predicted probability and low short interest are only 12% and 1% respectively; suggesting that high short interest for firms in column 2 is driven by arbitrage strategies. This finding is particularly striking as the short interest model (Table III) does not include either convertible securities outstanding or index membership as explanatory variables. Yet, consistent with hypothesis H3, the model has the ability to identify firms with characteristics related to arbitrage strategies.

We next examine firms with high predicted probability that have either low short interest (short sale constrained, column 3) or high short interest (valuation shorts, column 4). Consistent with the short-constraint hypothesis, firms with high predicted probability but low short interest are significantly smaller and have lower share turnover relative to firms with high predicted probability and high short interest. The fraction of firms that are S&P500 constituents or have convertible bonds outstanding is not economically different across two groups. Further, there is little economic difference in the BM ratio. Thus, although the model identifies a set of firms that subsequently experience significantly negative returns, the existence of short sale constraints likely prevents short sellers from assuming positions in these firms. These findings support the argument presented by Chen et al (2002) and Boehme et al (2005), among others, that the level of reported short interest is an imprecise proxy for identifying firms that are short constrained.

As a final additional test, we replicated the analysis in Panel A of Table VII 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 model have the ability to distinguish between valuation and arbitrage shorts over and above an indicator variable that captures the existence of convertible bonds? This analysis 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 arbitrage shorts would automatically classify short selling in almost 80% of the sample 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.[16] Specifically, we document that the firms with high short interest but low predicted probability from our model experience

[16] Detailed results are available from the authors upon request.

abnormal returns of 0.39% per month (not significant). In contrast, the firms with high short interest and high predicted probability experience abnormal returns of -0.68% 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.

V. Conclusions

This paper provides new insights into the decision process of information arbitrageurs. We examine a unique database of short sale recommendations that were motivated by a pessimistic opinion on firm valuation. We find that the firms in the short database experience significantly poor returns during the first year after the short recommendation was issued. Based on the short database, we build a parsimonious model that describes the decision process of short sellers. We find that the information contained in financial statement variables, such as accruals, gross margin, sales growth, and SG&A, and in valuation indicators, such as BM ratio and prior momentum, is related to the information set of short sellers. We designate firms with high predicted probability from this model as potential targets of valuation motivated short sellers. The *out-of-sample* tests strongly validate our model. We document that firms identified as valuation shorts exhibit high short interest and poor stock performance and that firms identified as arbitrage shorts exhibit a high fraction of S&P500 membership and convertible bonds.

Our key contributions are as follows. First, and most important, the study provides new insights into the information arbitrage process. We document that information arbitrageurs attempt to exploit the return predictability in valuation and fundamental signals, suggesting a direct link between the literature on the predictive ability of fundamental signals and the literature on the trading behavior of information arbitrageurs. Second, the study validates the importance of accounting-based information for short sellers by examining short recommendations. We argue that short recommendations allow for a more direct examination of the trading behavior of valuation-motivated short sellers than the level of short interest, which aggregates both valuation and arbitrage shorts. Third, we document that 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 and future returns in recent years.

References

Abarbanell, J.S., and B.J. Bushee, 1997, Fundamental Analysis, Future Earnings, and Stock Prices. Journal of Accounting Research Vol 35, number 1, 1997, pages 1-24

Abarbanell, J.S., and B.J. Bushee, 1998, Abnormal Returns to a Fundamental Analysis Strategy. The Accounting Review Vol 73, number 1, 1998, pages 19-45.

Agarwal, V., and N. Naik, 2005, Hedge funds, Foundations and Trends in Finance 1, 103-169.

Ali, A., L. Hwang and M. Trombley, 2003. Arbitrage Risk and the Book-to-market anomaly. Journal of Financial Economics 69 (2), 355-373

Arnold, T., A.W. Butler, T.F. Crack and Y. Zhang, 2005, The information content of short interest: A natural experiment, Journal of Business 78, 1307-1335.

Asquith, P., and L. Meulbroek, 1995, An empirical investigation of short interest, Working paper, MIT.

Asquith, P., P. Pathak, and J. Ritter, 2005, Short interest, institutional ownership, and stock returns, Journal of Financial Economics 78, 243-276.

Beneish, M. D., 1999, The detection of earnings manipulation, Financial Analysts Journal 55, 24-36.

Beneish, M.D., C. Lee, R. Tarpley, Contextual Fundamental Analysis Through the Prediction of Extreme Returns. Review of Accounting Studies Vol 6, Numbers 2-3, 2001, 165-189

Beneish, D and M. Vargus, 2002. Insider Trading, earnings quality and accrual mis-pricing. The Accounting Review, 77 (4), 755-792.

Boehme, R., B. Danielson, and S. Sorescu, 2005, Short sale constraints, dispersion of opinion and overvaluation, Journal of Financial and Quantitative Analysis, forthcoming.

Boehmer, E., C. Jones, and X. Zhang, 2005, Which shorts are informed?, Working paper, Texas A&M University.

Cao, B., D. Dhaliwal, and A. Kolasinski, 2007, Bears and Numbers: Investigating how short sellers exploit and affect earnings-based pricing anomalies, Working paper, University of Arizona.

Chen, J., H. Hong, and J. C Stein, 2002, Breadth of ownership and stock returns, Journal of Financial Economics 66, 171-205.

Christophe, S. E., M. G. Ferri, and J. J. Angel, 2004, Short-selling prior to earnings announcements, Journal of Finance 59, 1845-1875.

D'Avolio, G., 2002, The market for borrowing stock, Journal of Financial Economics 66, 271-306.

Dechow P. M., A. Hutton, L. Meulbroek, and R. Sloan, 2001, Short-sellers, fundamental analysis and stock returns, Journal of Financial Economics 61, 77-106.

Desai, H., K. Ramesh, S. R. Thiagarajan, and B. V. Balachandran, 2002, An investigation of the informational role of short interest in the Nasdaq market, Journal of Finance 57, 2263-2287.

Desai, H., S. Krishnamurthy, and K. Venkataraman, 2006, Do short-sellers target firms with poor earnings quality? Evidence from earnings restatements, Review of Accounting Studies 11, 71-90.

Diamond, D. W., and R. E. Verrecchia, 1987, Constraints on short-selling and asset price adjustment to private information, Journal of Financial Economics 18, 277–311.

Diether, K., K. Lee, and I. Werner, 2005, Can short sellers predict returns? Daily evidence, Working paper, Ohio State University.

Fama, E., and K. French, 1993, Common risk factors in the returns on stocks and bonds, Journal of Financial Economics 33, 3-56.

Fama, E. and K. French, 1995. Size and Book-to-market factors in earnings and returns. Journal of Finance 50 (1), 131-155.

Francis, J., M. Venkatachalam and Y. Zhang, 2005, Do short sellers convey information about changes in fundamentals or risk?, Working paper, Duke University.

Hirshleifer, D., S. Teoh, and J. Yu, 2005, Do short sellers arbitrage accounting-based anomalies?, Working paper, Ohio State University.

Hong, H., and J.C. Stein, 1999, A unified theory of underreaction, momentum trading, and overreaction in asset markets, Journal of Finance 54, 2143-2184.

Jegadeesh, N., and S. Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, Journal of Finance 48, 65-91.

Kraft, A., Leone, A., and C. Wasley. An Analysis of the Theories and Explanations Offered for the Mispricing of Accruals and Accrual Components. Journal of Accounting Research 44 (2006): 297-339.

Lamont, O., and C. Jones, 2002, Short sale constraints and stock returns, Journal of Financial Economics 66, 207-239.

Lev, B., and S. R. Thiagarajan, 1993, Fundamental information analysis, Journal of Accounting Research 31, 190-215.

Lakonishok, J., A. Shleifer, and R. Vishny, 1994, Contrarian investment, extrapolation, and risk, Journal of Finance 49, 1541-1578.

Mashruwala, C. S. Rajgopal and T. Shevlin. 2006. Why is Accruals anomaly not arbitraged away?, Journal of Accounting and Economics 21, 107-138.

Miller, E. M., 1977, Risk, uncertainty and divergence of opinion, Journal of Finance 32, 1151-1168.

Mitchell, M., T. Pulvino, and E. Stafford, 2004, Price pressure around mergers, Journal of Finance 59, 31-63.

Nagel, S., 2005, Short sales, institutional investors, and the cross-section of stock returns, Journal of Financial Economics 78, 277-309.

Piotroski, J.D., 2000, Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers. Journal of Accounting Research Vol 38, Supplement, 2000, pages 1-41

Richardson, S. A., R. G. Sloan, M. T. Soliman, and I. Tuna, 2005, Accrual reliability, earnings persistence and stock prices, Journal of Accounting and Economics 39, 437-486.

Schipper, K., and L. Vincent, 2003, Earnings quality, Accounting Horizons 17 (supplement), 97-110.

Senchack, A.J., and L.T. Starks, 1993, Short sale restrictions and market reaction to short interest announcements, Journal of Financial and Quantitative Analysis 28, 177-194.

Sias, R., and L.T. Starks, 1997, Institutions, individuals and the turn-of-the year, Journal of Finance 52, 1543-1562.

Sloan, R. G., 1996, Do stock prices fully reflect information in accruals and cash flows about future earnings?, The Accounting Review 71, 289-315.

Table I
Performance of firms in the short database

The table presents the performance statistics for the sample of firms identified by an independent research firm as potential targets for short selling. The sample comprises all 67 firms that were targeted for short selling by the firm from September 1998, until June 2005. Panel A lists the mean (median) raw buy and hold return and the market-adjusted buy and hold return for year -1, the event month, and year +1. Panel B reports the types of 'bad news' events that were reported in year +1.

Panel A: Returns

	Months (-12, -1)	Month 0	Months (+1, +12)
Raw returns (%)	77.90 ***	-4.03 **	-9.71
	(24.55) ***	(-4.03) *	(-18.32) **
Market-adjusted returns (%)	72.36 ***	-4.89 **	-15.02 **
	(31.85) ***	(-3.57) **	(-17.12) ***

Panel B: Summary of 'bad news' events in year +1

Events		# Observations
Firms with 'bad news' reported		42
Reported lower than expected earnings	24	
Lowered guidance on future earnings / sales	20	
Analyst downgrade	23	
Earnings restatement, earnings or audit delay, accounting concerns	7	
Regulatory action, Lawsuit, SEC investigation	8	
Top management turnover	2	
Ratings cut, ratings watch, covenant violation	3	
Firms with no 'bad news' reported		25

*, ** and *** denote significance at the 10-, 5- and 1-percent level respectively.

Table II
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 variable definitions are reported in Appendix 1.

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

*, ** and *** denote significance at the 10-, 5- and 1-percent level respectively.

Table III
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 to target a firm. 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.

Dependent variable: '1' for target, '0' otherwise				
	Unadjusted, winsorized		Industry-median adjusted, winsorized	
Model	(1)	(2)	(3)	(4)
Intercept	-5.7745 ***	-5.9375 ***	-6.9265 ***	-6.9941 ***
Performance variables				
BM	-2.3495 ***	-2.0908 ***	-2.1931 ***	-2.1489 ***
Prior 1-year return	0.2308 *	0.2289 *	0.2768 **	0.2662 **
Financial variables				
DSRI	-0.0200	-0.0210	0.0128	0.0162
GMI	-0.2040	-0.2286 *	-0.2205 *	-0.2397 *
SGAI	0.7913 **	0.6395 *	0.8383 **	0.6644 *
AQI	-0.2267	-0.0318	-0.2929	-0.0809
SGI	0.2648 ***	0.1494 *	0.5047 ***	0.2910 **
DEPI	-0.2518	-0.4099	-0.2141	-0.4101
LVGI	-0.1991	-0.1406	-0.2058	-0.1595
TOTACC		2.0702 ***		2.3394 ***
OPACC	2.8047 ***		3.0838 ***	
Firm characteristics				
Size	0.0000	0.0000	0.0000	0.0000
Turnover	47.6647 *	39.1111	57.0076 **	49.3988 *
Pseudo R^2 (%)	9.2	10.2	10.1	11.3

*, ** and *** denote significance at the 10-, 5- and 1-percent level respectively.

Table IV
The level of short interest
categorized by the predicted probability from the short interest model

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 III. The decile portfolios are assigned 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 III 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.

Short interest (%)				
	Unadjusted		Industry-median adjusted	
From Table II	Model (1)	Model (2)	Model (3)	Model (4)
Decile 1 (Low)	0.510	0.518	0.465	0.480
2	0.666	0.605	0.634	0.618
3	0.792	0.807	0.810	0.793
4	0.872	0.898	0.850	0.911
5	0.966	0.932	1.044	0.956
6	1.077	0.989	1.055	1.049
7	1.134	1.240	1.193	1.243
8	1.545	1.376	1.590	1.441
9	1.947	1.948	1.960	2.030
Decile 10 (High)	3.312	3.466	3.239	3.274
t for High-Low	24.48	25.40	24.75	24.88

Table V
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. 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 III. The decile portfolios are assigned 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 III 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.

Abnormal returns

	Unadjusted		Industry-median adjusted	
From Table II	Model (1)	Model (2)	Model (3)	Model (4)
Decile 1 (Low)	0.0111 ***	0.0104 ***	0.0125 ***	0.0118 ***
2	0.0074 ***	0.0073 ***	0.0073 **	0.0073 **
3	0.0015	0.0031	0.0050 **	0.0069 ***
4	0.0049 **	0.0043 **	0.0033	0.0020
5	0.0007	0.0012	0.0014	0.0024
6	-0.0007	0.0012	0.0002	-0.0015
7	0.0004	-0.0002	-0.0004	-0.0003
8	-0.0002	-0.0024	-0.0013	-0.0013
9	-0.0042 *	-0.0044 **	-0.0055 ***	-0.0059 ***
Decile 10 (High)	-0.0074 **	-0.0059 *	-0.0076 **	-0.0064 **

*, ** and *** denote significance at the 10-, 5- and 1-percent level respectively.

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 based on the predicted likelihood of being targeted by informed short sellers. We estimate three alternative specifications of the logistic regression model, similar to those in Table III – 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 assigned 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.

Abnormal returns

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

*, ** 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 the 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 (average short interest, %) [number of observations] 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

Predicted probability	Low (deciles 1 - 3)	Medium (deciles 4 - 7)	High (deciles 8 - 10)
Low short interest (decile 1)	1.19 *** (0.03) [1320]	0.14 (0.04) [900]	-1.08 ** (0.04) [435]
High short interest (decile 10)	0.08 (7.43) [264]	-0.72 *** (7.26) [623]	-0.71 *** (8.74) [1237]

Panel B: Firm characteristics

Predicted probability	Low (deciles 1-3)		High (deciles 8-10)	
Short interest	Low (decile 1)	High (decile 10)	Low (decile 1)	High (decile 10)
MVE ($ mill)	48.6	726.9 ###	229.0	926.8 ###
BM	1.55	1.22 ###	0.32	0.27 ###
Average turnover (%)	0.13	0.42 ###	0.25	0.76 ###
% with cvt. sec.	12.3	56.4 ###	19.5	24.0 #
% in S&P 500 Index	0.53	19.7 ###	0.23	5.01 ###

*, ** 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.

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-1' through September of year 't'
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)	$[(\#12_{t-1} - \#41_{t-1}) / \#12_{t-1}] / [(\#12_t - \#41_t) / \#12_t]$
SGAI (Selling, general, and admn. expenses index)	$[\#189_t / \#12_t] / [\#189_{t-1} / \#12_{t-1}]$
AQI (Asset quality index)	$[(1-(\#4_t + \#8_t)) / \#6_t] / [(1-(\#4_{t-1} + \#8_{t-1})) / \#6_{t-1}]$
SGI (Sales growth index)	$\#12_t / \#12_{t-1}$
DEPI (Depreciation index)	$[(\#14_{t-1} - \#65_{t-1}) / (\#14_{t-1} - \#65_{t-1} + \#8_{t-1})] / [(\#14_t - \#65_t) / (\#14_t - \#65_t + \#8_t)]$
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-1' to September of year 't'.
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'.