Short interest, returns, and fundamentals

Ferhat Akbas[†], Ekkehart Boehmer[#], Bilal Erturk[◊], and Sorin Sorescu*

August 31, 2013

Abstract

We show that short interest predicts future bad news, negative earnings surprises, and downward revisions in analyst earnings forecasts. Moreover, short interest is a better predictor of changes in firm fundamentals for stocks that are harder to short and short sellers appear to have information about these events several months before they become public. Most importantly, the well-known cross-sectional relation between short interest and future stock returns vanishes after controlling for short sellers' information about future fundamental news. Thus, short sellers contribute in a significant manner to price discovery about firm fundamentals.

Keywords: Short interest, fundamental information, cross-section of stock returns
JEL Classification: G12, G14

[†] School of Business, University of Kansas, Lawrence, KS 66045-7601. [#]EDHEC Business School, One George Street, 049145 Singapore. [◊] Spears School of Business, Department of Finance, Oklahoma State University, Stillwater, OK 74078-4011. * Mays Business School, Department of Finance, Texas A&M University, College Station, TX 77843-4218. We are indebted to Bart Danielsen for providing short interest data. We thank Kerry Back, Dave Blackwell, Mike Gallmeyer, Shane Johnson, Dmitry Livdan, and Jay Ritter for useful comments.

Short interest, returns, and fundamentals

Abstract

We show that short interest predicts future bad news, negative earnings surprises, and downward revisions in analyst earnings forecasts. Moreover, short interest is a better predictor of changes in firm fundamentals for stocks that are harder to short and short sellers appear to have information about these events several months before they become public. Most importantly, the well-known cross-sectional relation between short interest and future stock returns vanishes after controlling for short sellers' information about future fundamental news. Thus, short sellers contribute in a significant manner to price discovery about firm fundamentals.

Keywords: Short interest, fundamental information, cross-section of stock returns
JEL Classification: G12, G14

1. Introduction

One robust finding in the empirical investments literature is the negative cross-sectional relation between the level of short interest and future abnormal stock returns. Stocks that are heavily shorted have dismally low returns over the following month. This suggests that short sellers are informed traders, but it is less clear what kind of information motivates their trades. In this paper, we contribute to this literature and shed light on the nature of short sellers' information.

Anecdotal evidence from professional short sellers provides hints about how short sellers become informed. These traders view their role akin to that of corporate "detectives," who seek to uncover and expose problems with publicly traded firms. Famed short seller James Chanos of Kynicos Associates illustrated this view in testimony before the US Congress:

> Finally, I want to remind you that, despite two hundred years of "bad press" on Wall Street, it was those "un-American, unpatriotic" short sellers that did so much to uncover the disaster at Enron and at other infamous financial disasters during the past decade (Sunbeam, Boston Chicken, etc.). While short sellers probably will never be popular on Wall Street, they often are the ones wearing the white hats when it comes to looking for and identifying the bad guys! *James Chanos, Testimony before the House Committee on Energy and Commerce (February 6, 2002).*

The argument above suggests that short sellers' trades contain negative information about future changes in firm fundamental value. If limits to arbitrage prevent instantaneous stock price adjustments, then short sellers' unfavorable information on future fundamentals represents one plausible explanation for the well-known negative cross-sectional relation between short interest and future stock returns.

We present novel evidence that short sellers' positions predict deterioration in future firm fundamentals in a way that is consistent with the return predictability relation observed in prior studies. Combining short interest data with RavenPack's dataset of all public news' content analysis during 2000-2010, we show that that short sellers take profitable positions in stocks that will experience negative public news in the future. Likewise, short sellers seem to avoid stocks with positive future public news. Short sellers also correctly anticipate the outcome of future earnings surprises and future revisions in analyst earnings forecasts.

Our results suggest that the cross-sectional relation between short interest and future stock returns can be explained, in large part, by the information that short sellers have about future changes in firm fundamentals. Moreover, short interest is a better predictor of changes in firm fundamentals for stocks that are harder to short, as predicted by Diamond and Verrecchia's (1987) argument that shorting costs first remove the less informed traders from the market.

We make two important contributions to the literature. First, we show that short sellers are informed about future changes in firm fundamentals. Second, we show that after accounting for this information, the well-known predictive relation between short interest and future return vanishes. Thus, short-sellers join the ranks of stock analysts, institutional investors, underwriters, auditors, and bank lenders, who have been shown to provide similar information-acquiring functions.[1]

[1] For example, underwriter quality has been shown to influence the extent of IPO underpricing (Beatty and Ritter, 1986; Carter and Manaster, 1990), as has the identity of an IPO firm's auditors (Beatty, 1989). A long literature shows the positive effect of bank loan announcements on a firm's stock price (e.g., Mikkelson and Partch, 1986; James, 1987; Lummer and McConnell, 1989), including the finding that announced loans from higher-quality lenders are associated with more positive borrower abnormal returns (Billett, Flannery, and Garfinkel, 1995). Brennan and Subramanyam (1995) report that the equity of firms that are followed by a larger number of investment analysts trade with smaller bid-

Our findings are important in view of concerns raised by regulators about other – potentially more nefarious – consequences of shorting activities. For example, in the midst of the 2008 financial crisis, short sellers were regularly suspected to have intentionally driven stock prices below their fundamental values. In line with these views, the Securities and Exchange Commission (SEC) temporarily banned the shorting of financial institutions. SEC justified the action by claiming that "unbridled short selling is contributing to the recent, sudden price declines in the securities of financial institutions unrelated to true price valuation" (SEC release 2008-211).[2]

Although our results provide strong evidence that short sellers' trades predict returns due to their informativeness about future fundamentals, we cannot completely rule out that some short trades are manipulative or that some short sellers have manipulative intentions. Indeed, some studies have shown instances in which the trading patterns of some short sellers are consistent with a manipulative intent.[3] However, our results do suggest that – on average – the dominant effect of short positions is to *predict*, rather than *cause*, negative abnormal stock returns. This is because short positions are informative about value-relevant events that are exogenous and cannot be systematically manipulated by traders: news releases, earnings surprises, and analyst forecast revisions.

Our paper builds upon a broadening base of empirical research demonstrating that short sellers are informed traders who predict future returns (see e.g., Desai et al., 2002;

---

ask spreads, reflecting lower informational asymmetries across traders in the market. Boehmer and Kelley (2009) report that institutional investors improve the efficiency of security prices.

[2] In contrast with SEC's assertion, Boehmer, Jones, and Zhang (2013) show that this ban imposed significant costs on traders in terms of deteriorating liquidity and higher volatility, without any permanent effects on share valuations.

[3] See, e.g. Henry and Koski (2010).

Boehmer, Jones, and Zhang, 2008; Diether, Lee, and Werner, 2009; Asquith, Pathak, and Ritter, 2005) and facilitate price discovery (see e.g., Saffi and Sigurdsson, 2012; Boehmer and Wu, 2013). The literature also shows that short sellers typically initiate short positions when they can infer low fundamental values from public sources. For example, short sellers might: (i) engage in forensic accounting, looking for high levels of accrual as evidence of hidden bad news (Hirshleifer, Teoh, and Yu, 2011), (ii) detect financial misconduct (Karpoff and Lou, 2010), (iii) target firms with poor earnings quality (Desai, Krishnamurthy and Venkataraman, 2006), or (iv) look for abnormally elevated price-earnings ratios (Dechow et. al, 2001). In some cases, short sellers have been shown to increase short positions just before announcements of unfavorable news affecting the underlying company (Christophe, Ferri, and Angel, 2004; Christophe, Ferri, and Hsieh, 2010).

Closely related to our paper is the analysis in Engelberg, Reed, and Ringgenberg (2012), who examine daily short selling flows over shorter, ten-day periods preceding news announcements. They find that the highest shorting volume in their sample occurs during the day of the news announcement rather than the preceding trading days. They conclude that short sellers' returns are related to their ability to analyze, rather than predict, public news announcements, and that short sellers are highly skilled in interpreting public news announcements.

Our paper highlights a different, but equally important skill set of professional short-sellers: the ability to predict, several months ahead, changes in firm fundamental values. Using a different experimental design from Engelberg, Reed, and Ringgenberg (2012), we show that the most heavily shorted stocks are followed by events associated with a decrease in firm fundamental value, which occur as late as twelve months into the future.

An implicit assumption in Engelberg, Reed, and Ringgenberg's (2012) study is that short sellers have short investment horizons that do not extend beyond two calendar weeks. While this can be true for some traders, casual observations of professional short sellers suggest that their investment horizon is much longer. For example, James Chanos started to short Enron's stock in October 2000, months before the company's demise in 2001. Other professional short sellers whose shorting strategies are based on "forensic accounting" (such as shorting firms with high accruals), typically adopt investment horizons ranging from several months to several years. In our analysis we allow for such longer horizons. Our results suggest that this methodological variation matters—over this longer period, short sellers do seem able to predict news announcements as well as changes in firm fundamental value.

Our results are complementary to the findings in Englelberg, Reed, and Ringgenberg (2012). The fact that short sellers are able to predict changes in firm fundamentals several months into the future does not preclude them from also being able to interpret public news on the day of the news release. Our paper, therefore, provides a more complete picture of the role played by short sellers in financial markets: we demonstrate that short sellers are not only short-term skilled information processors, but also generators of long-lived value-relevant information about fundamental value. The duration or "shelf life" of short seller's information set could be as long as twelve months, significantly longer than previously documented in the literature.

Our results are supportive of Diamond and Verrecchia's (1987) prediction that shorts are more informed in the presence of high shorting costs. Using various proxies for short sale constraints, we find that short interest better predicts changes in firm fundamentals for stocks that are more difficult to short.

Finally, our paper does not preclude alternative explanations of the negative relation between short interest and subsequent stock returns. Several empirical studies (see e.g. Asquith, Pathak, and Ritter (2005)) provide evidence consistent with Miller's (1977) argument that short-sale constraints cause stock prices to be overvalued because investors with negative valuations cannot easily impact prices through trading. These studies argue that stocks that are already heavily shorted are the most difficult to short. Our paper shows that in addition to being a proxy for short sale constraints, the level of short interest (which shows positions already established by short sellers) is also informative about future changes in firm fundamental value, thus providing another transmission channel for the relation between short interest and future stock returns.

The rest of the paper is organized as follows. In Section 2, we describe the data sources, methodology, and variables used as proxies for firm fundamental information. We present empirical results in Section 3 regarding the informational content of short interest and its role on return predictability. Section 4 shows short interests' return predictability vanishes after controlling for short's informational content. Section 5 presents robustness checks, and Section 6 concludes.

2. Data and methodology

We used two different methods to assess the extent to which short interest predicts stock returns or changes in firm fundamentals. First, we use monthly Fama-MacBeth regressions (Fama and French, 1992; Carhart, 1997), where we regress future stock returns or changes in firm fundamentals on current short interest levels. To measure the economic significance of the results, we replicate the cross-sectional Fama-MacBeth regressions after transforming all independent variables into decile ranks and then standardize these ranks to take values between zero and one. This transformation makes interpreting and comparing coefficients across variables more intuitive. We refer to these standardized regressions in the paper but tabulate the full results in the Internet Appendix to this paper.

The second method is the calendar time portfolio procedure (Jegadeesh and Titman, 1993). Each month, we rank stocks on the basis of short interest around its mid-month release date and assign them to decile portfolios. For each portfolio we compute subsequent changes in firm fundamentals, beginning with the calendar month after the short interest release. Alternatively, we compute equally-weighted raw and abnormal returns for each portfolio. Abnormal returns are computed as the intercept from the standard Fama-French-Carhart four-factor model:

$$r_{pt} - r_{ft} = \alpha + \beta_1 (r_{mt} - r_{ft}) + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 MOM_t + \varepsilon_{pt} \quad (1)$$

2.1. Level of short interest

Our sample includes all common stocks listed on NYSE, Amex, and NASDAQ (share code 10 or 11), between January 1988 and December 2010, that are included in the daily CRSP (*Center for Research on Security Prices* at the University of Chicago) data.

We measure monthly levels of short interest, *Short*, using the levels of outstanding shares sold short on NYSE, Amex, and Nasdaq divided by the number of shares outstanding. The number of shares outstanding and daily returns are obtained from CRSP.

2.2. Proxy for shorting costs

For our main tests we use institutional ownership (*IO*) as a measure of shorting costs because it is correlated with the supply of lendable shares (D'Avolio, 2002; Asquith, Pathak, and Ritter, 2005; Nagel, 2005).[4] Data on institutional ownership are obtained from 13-F filings, available from Thomson Financial. We define institutional ownership, *IO*, as the sum of the holdings of all institutions for each stock in each quarter, divided by the number of shares outstanding. Stocks that have available return data but no reported institutional holdings are assumed to have zero institutional ownership.

Institutional ownership is highly correlated with firm size. To address this problem, we orthogonalize it with respect to firm *size* (market value of equity). Since institutional ownership is bounded between 0 and 1, each quarter we group stocks into size deciles and run the following cross-sectional regression on a quarterly basis:

$$IO_{jt} = \sum_{i=1}^{10} \alpha_{it} Size_{ijt} + \varepsilon_{jt} \tag{2}$$

The explanatory variables are ten dummy variables that represent the ten size-based portfolios. For instance if firm j is in size decile *1* in quarter t, then $Size_{1jt}$ equals 1 and all other dummy variables equal zero. By construction, the intercept term is omitted from the regressions. We refer to the residuals from the quarterly cross-sectional regressions as residual institutional ownership (*RIO*). *RIO* measures the deviation of a firm's institutional ownership from the average institutional ownership within its size decile, in any given quarter.

---

[4] In the robustness section we explore other proxies for short selling costs, such as market capitalization (SIZE), Amihud's (2002) measure of illiquidity, share turnover, and analyst coverage. The results are similar to those obtained with the main (IO) proxy.

2.3. Stock returns and excess stock returns

Our analysis uses both raw and abnormal stock returns. To compute abnormal returns, we risk-adjust monthly returns (from CRSP) using size, momentum, and book-to-market factors (Fama and French, 1992; Jegadeesh and Titman, 1993; Carhart, 1997). For the Fama-MacBeth (1973) regression method, we use monthly levels of size, momentum, and book-to-market for each stock in our data set. As before, *size* is the market value of equity obtained from CRSP. Book-to-market (*B/M*) is the ratio of book value of equity to market value of equity. Book value is computed as in Fama and French (2002) and is measured at the most recent fiscal year-end that precedes the calculation date of market value by at least three months. We exclude firms with negative book values. The book value data are obtained from COMPUSTAT. The stock's return momentum (*R_MOM*) is the cumulated raw return over the previous 12-month period. For the calendar time portfolio approach, we obtain from Kenneth French's website at Dartmouth College the standard four-factor series: R_{mt}-R_{ft} (the market excess return), *SMB* (the size factor return), *HML* (the book-to-market factor return), and *MOM* (the momentum factor return).

2.4. Proxies for future changes in firm fundamentals

We examine three types of proxies for future changes in firm fundamentals: (i) future public news, (ii) earnings surprises, and (iii) changes in analyst forecasts.

2.4.1. Public news

We obtain data on public news from RavenPack News Analytics, edition 3.0. RavenPack is an established provider of news analysis services. Since January 1, 2000, they collect corporate news items from all public sources, including Dow Jones, Barron's, and the Wall Street Journal, and classify each news item through content analysis according to its sentiment, relevance, topic, novelty, and market impact. The most relevant category for our paper is RavenPack's *Event Sentiment Score (ESS)* variable*.* This variable measures whether a particular news item contains "good" or "bad" information about the underlying corporation. The *ESS* variable takes values from 0 to 100, with lower scores indicating bad news and higher scores indicating good news. A score of 50 is classified as neutral.[5]

News items in RavenPack's database are based on public news sources that sometimes mention more than one corporation. If the original news source mentions two or more corporations, RavenPack creates a separate news item for each corporation. However, RavenPack uses the "relevance" variable to differentiate between news items where the corporation is the main object of the original news source and news items where the name of the corporation is mentioned only tangentially. The relevance variable takes values from 0 to 100. Throughout our analysis, we focus on news items where the sample firm's name is the main object of the news story and eliminate all news items with a relevance score lower than 50.

RavenPack assigns a unique corporate identifier to each stock in their database that we manually match to CRSP's PERMNO security identifier. We successfully match 5,310 securities from CRSP with RavenPack data. If, for any of these securities, RavenPack

[5] RavenPack calculates sentiment scores according to an algorithm that is reportedly capable of "interpreting actual figures, estimates, ratings, revisions, magnitudes and recommendations disclosed in news stories." RavenPack states that their algorithm can also "compare actual vs. estimated figures about earnings, revenues, and dividends – and produce an ESS score based on comparisons… (page 12)." Additional details on Event Sentiment Scores are contained in RP News Analytics – Dow Jones Edition – User Guide v.3.0, available at www.ravenpack.com. See Appendices B and D of that document for more details on ESS scores.

contains no news items in a given month, we assume no news about that particular stock that month. The remaining stocks are not covered by the RavenPack data.

We construct two monthly news measures using RavenPack's event sentiment score (ESS). The first measure, *NEWS1*, is constructed using a trinary classification of news content. For each event we assign a numerical value of either +1, 0, or –1, depending upon the level of the ESS score. Values of +1 are assigned to news items with positive sentiment (ESS score equal to 51 or higher). Values of 0 are assigned to news items having neutral sentiment (ESS score equal to 50), and values of –1 are assigned to news items with negative sentiment (ESS score equal to 49 or lower). If a stock does not have a news item in a particular month, we assume that this is equivalent to having neutral news content, and assign a value of zero to the trinary variable during that month.[6] Each day, we average these trinary variables for all news events in RavenPack, resulting in a daily news content measure. *NEWS1* is then computed as the monthly average over these daily news content measures.

For our second measure, *NEWS2*, we use the actual ESS score assigned by RavenPack to the news event. We follow a similar averaging process as we did in constructing *NEWS1*, but instead of using trinary variables, we use the actual ESS values (ranging from 0 to 100). If a stock does not have an ESS score for a particular month, we treat that month as "neutral" and assign a value of 50 to NEWS2.

[6] In additional untabulated robustness checks, we use different ESS cut-off values for the definition of "good," "neutral," and "bad" news. For example, instead of using the values of 51 and 49 as cutoffs for ESS scores, we use (i) 75 and 25, (ii) 66 and 33, and (iii) 80 and 20. The results are similar to those reported in the paper. We also repeat our analysis by excluding observations without identifiable news. Again, the results are similar.

We perform three additional modifications to enhance the reliability and integrity of both *NEWS* measures. First, we eliminate securities with prices below $5 at the end of the month for which short interest is measured. Second, to mitigate concerns about reverse causality between stock prices and news, we eliminate all news whose contents are categorized as "stock prices" or "order imbalances." [7] Third, we recognize that bigger firms have more media coverage and are also more likely to work with investor relations companies to help frame the news items in a more positive light (Solomon, 2012). To account for these potential biases, we orthogonalize *NEWS1* and *NEWS2* with respect to firm *size*, following the procedure we use for institutional ownership.

2.4.2. Earnings surprises

We construct earnings surprises from quarterly earnings announcements from COMPUSTAT. We use two measures of earnings surprises: standardized unexpected earnings (*SUE*) and cumulative abnormal returns around the earnings announcements (*CAR*).

Following Foster, Olsen, and Shevlin (1984) and Chan, Jegadeesh, and Lakonishok (1996), we define *SUE* in quarter *q* as

$$SUE_q = \frac{EPS_q - E[EPS_q]}{\sigma_q} \qquad (3)$$

where q is the quarter, EPS_q are the most recent quarterly earnings per share, $E[EPS_q]$ are expected earnings per share, and σ_q is the standard deviation of unexpected earnings (EPS_q –

[7] The remaining 26 news categories are Acquisitions-Mergers, Analyst-Ratings, Assets, Bankruptcy, Corporate-Responsibility, Credit, Credit Ratings, Dividend, Earnings, Equity-Actions (i.e., buybacks), Exploration, Index (i.e., index delisting), Industrial-Accident, Insider-Trading, Investor-Relation, Labor-Issues, Legal Issues, Marketing Partnerships, Price-Targets, Products-Services, Regulatory, Revenues, Taxes, Transportation.

$E[EPS_q]$) over the preceding eight quarters. To estimate expected earnings we use a seasonal random walk model as in Chan, Jegadeesh, and Lakonishok (1996):

$$E[EPS_q] = \alpha + EPS_{q-4} \tag{4}$$

The second measure of earnings surprises is the daily cumulative market-adjusted return around the earnings announcement date (*CAR*). *CAR*s are computed as in Brown and Warner (1985):

$$CAR_i = \sum_{j=-2}^{j=2} (r_{ij} - r_{mj}) \tag{5}$$

where r_{ij} is the stock i's return on day j and r_{mj} is the return on the equally-weighted CRSP market index on day j. Day j=0 is the earnings announcement date.

2.4.3. Changes in analyst forecasts

We measure changes in analysts' consensus earnings-per-share forecasts for the current fiscal year-end using data from FIRST CALL. We begin our sample period in 1990 because FIRST CALL data are incomplete for previous years. We use the latest mean estimate available in a given month as our consensus forecast and measure the change in consensus estimate as the difference in mean estimates from the previous month ($\Delta EPS_t = EPS_t - EPS_{t-1}$). We normalize ΔEPS_t in two ways. First, we divide it by the absolute value of the consensus forecast at the end of the previous month ($|EPS_{t-1}|$). Alternatively, we divide it by the stock price at the end of the previous month (P_{t-1}).

3. Results

We begin by replicating the negative predictive relation between the level of short interest and the cross-section of stock returns. We then examine the source of this predictability by investigating the relation between short interest and future public news, earning surprises, and changes in analyst forecasts.

3.1. Short interest and future returns

Table 1 presents results on *Short's* ability to predict stock returns, using calendar time portfolios. As expected, heavily shorted stocks significantly underperform lightly shorted stocks. The difference is 0.73% per month on an absolute basis, and by 0.95% per month on a risk-adjusted basis. For the most heavily shorted stocks (decile 10), the average relative short interest is 11.79% and the average monthly risk adjusted return is -0.57% (t=-3.23). These results corroborate previous studies, but do not inform on the cause of this relation.

3.2. Short interest and future fundamental news

We now examine the relation between short interest and future public news. If higher levels of short interest contain negative information about firm fundamentals, stocks with high short interest should predict less favorable news.

We begin with the calendar time portfolio methodology. The sample period for these tests is 2000-2010, corresponding to the availability of the RavenPack news dataset. The sample consists of 267,971 stock-month observations, 137,535 of which are from NASDAQ and 130,436 are from NYSE/AMEX. On average, our sample contains 2,042 stocks each month, 1,032 from NASDAQ and 1001 from NYSE/AMEX. Panel A of Table 2 presents key descriptive statistics for this sample. The average firm size is 4.92 billion dollars and exhibits considerable variation. On average, news content is neutral (zero, because we orthogonalize news with respect to size, resulting in a mean-zero residual) and average relative short interest is 4.88%.

We next examine the reliability of the news classification in the RavenPack data. If RavenPack categorization correctly captures the extent to which a news item surprises the market and conveys fundamental information about firms, we expect to find abnormal returns on news days compared to days without news or days with neutral news. We also expect these abnormal returns to be correlated with RavenPack's sentiment (ESS) variable that carries information about news content.

To test this prediction, we sort stocks each day into one of four groups (no news, bad news, neutral news, good news) based upon the daily averages of the NEWS1 and NEWS2 variables previously described. We compute daily mean and median excess return values for each group. We then calculate time series averages of these cross sectional mean and median values.

The results are presented in Panel B of Table 2. The top line in the panel shows time series mean excess returns for stocks for various news content groups. The mean daily excess return for stocks without RavenPack news is 6.9 basis points per day. The mean daily excess returns are 2.8 and -0.4 basis points per day for neutral news, using NEWS1 and NEWS, respectively. For news event that are categorized as "bad" by RavenPack, the mean daily excess return is -141.8 basis points for NEWS1 and -142.3 basis points for NEWS2. For news events classified as "good" by RavenPack, the mean daily excess returns are 109.5 and 131.8 basis points for NEWS1 and NEWS2, respectively. Similar results are found with median return values in the bottom line of Panel B.

The results in Panel B validate RavenPack's classification of news content along the ESS variable. Excess returns are significantly positive for news items classified as "good" by the ESS variable, and significantly negative for news items classified as "bad." These results also validate our methodological choice of assigning "no news content" to days when RavenPack news are absent: excess returns for "no news days" are close to zero and very similar to those measured on days when the ESS score is "neutral." Overall, these finding confirm that the RavenPack classifications are proper and that they contain actual unexpected news.

If short sellers are detectives who uncover unfavorable information, bad news should follow periods of heavy shorting. The results in Panel C of Table 2 readily confirm this conjecture. In a given month, the most heavily shorted stocks have the most negative public news the following month. At the other extreme, the least shorted stocks have the most positive public news the following month. The relation between short interest and future news is almost monotonically decreasing, for both the *NEWS1* and *NEWS2* measures. To assess the statistical significance of this relation we compute the difference in news content between stocks belonging to the top and bottom short interest decile. This difference, shown at the bottom of the panel, is strongly significant at the 1% level for both news measures.

Panel C also shows the univariate relation between short interest and future stock returns, for the purpose of validating that the relation remains negative in this sample that is constrained by the availability of RavenPack news data. These sorts provide a first indication that short interest predicts both news and return.

Diamond and Verrecchia (1987) argue that in the presence of non-trivial shorting costs, only investors who have strong negative information will choose to short. The heightened cost of shorting will discourage less informed investors from shorting. Thus, the level of short interest should be more informative in the presence of high shorting costs. We test this prediction using *residual institutional ownership* (*RIO*) as a proxy for shorting costs.[8] Diamond and Verrecchia's model implies that the relation between short interest and future public news is stronger for stocks with low *RIO*.

We sort stocks independently into terciles based on the level of *RIO* and into quintiles based on the level of short interest. We compute next month's average news content for each of the resulting 15 portfolios, and then average across all months. Panel D of Table 2 shows the conditional relation between short interest and future news, as a function of institutional ownership. As expected, the predictive ability of short interest (measured as the difference in news content between the top and bottom short interest quintiles) is a decreasing function of residual institutional ownership, and the conditional effect of institutional ownership is significant at better than the 1% level.

Overall, the results presented in Table 2 provide evidence that short interest predicts public news. Consistent with Diamond and Verrecchia (1987), this result is stronger when short sale costs are higher. Our results also suggest that both high and low levels of short interest carry informational content. High short levels predict negative news and low short levels predict positive news. Consistent with Boehmer, Huszar, and Jordan (2010), predictability is stronger for low short interest portfolios.

An advantage of the calendar time portfolio method is that it can detect non-monotonic relations between variables, but a possible shortcoming is that correlations with

[8] Additional proxies for short sale constraints are examined in the robustness section of the paper.

excluded variables could result in spurious relations. To address this concern, we estimate Fama-MacBeth regressions that control for residual institutional ownership, lagged news content (news content measured in month t-1), size, book-to-market, and momentum and present the results in Table 3. Consistent with the portfolio sorts, short interest continues to predict subsequent public news content. The coefficient on *short* is significantly negative for both the *NEWS1* and *NEWS2* measures and the relation is stronger for lower levels of residual institutional ownership (the difference between these coefficients (not tabulated) is statistically significant at the 1% level).

In the Appendix (Table A1) we report regressions based on standardized decile ranks. Using the same controls as in Table 2 Panel C, the difference between *NEWS1* in the highest short interest decile and *NEWS1* in the lowest short interest decile is -0.07 (the coefficient of *Short* in the first column of Table A1). This is similar in magnitude to the *NEWS1* spread of -0.064 shown in Panel B of Table 2. In addition, consistent with the results in Panel C of Table 2, we find that that the coefficient of *Short* increases for lower values of residual institutional ownership. The results are similar for the *NEWS2* measure.

Overall, the results presented in Table 3 corroborate our previous findings that short interest predicts future news. Our results also suggest that short sellers are not merely taking positions based on current news; rather, they predict future news. Short interest contains information about future news content even after controlling for current news content. Therefore, we rule out the possibility that our results are simply driven by momentum in news items and short sellers' ability to better process news items ex-post and take position accordingly. In fact, variation in current short interest predicts more variation in future news than it predicts variation in future returns.

3.3. Short interest and future earnings surprises

We now sharpen the focus of our analysis and look specifically at earnings announcements. Panel A of Table 4 reports descriptive statistics of portfolio sorts on earnings surprises. Specifically, we sort stocks into quintiles based on standardized unexpected earnings (*SUE*). For each group, we report the average value of *SUE*, as well as the average cumulative abnormal returns (*CARs*) surrounding earning announcements. As expected, the two measures are highly correlated. Stocks with the most positive earnings surprises earn *CARs* of +1.55% during the five-day announcement window. At the other extreme, *CARs* are –1.65% for stocks with the most negative earnings surprises.

In Panel B we sort stocks into quarterly short interest quintiles and compute the next quarter's earnings surprises (*SUE* and *CARs*) for each quintile. These are then averaged intertemporally across all quarters. Short interest predicts future earnings surprises for both the *CAR*s and *SUE* measures, and the predictability is statistically significant at the 1% level in both cases. These results corroborate previous evidence that short interest is informative about changes in firm fundamentals that are yet to be incorporated into stock prices. We explore the incremental role of shorting costs in Panels C and D. The predictive power of *Short* for *SUE* (Panel C) and *CAR* (Panel D) is greater for lower levels of residual institutional ownership (*RIO*), suggesting that existing short positions are more informative when constraints are more binding.

Table 5 presents cross-sectional regressions that examine short interest's ability to predict future earnings surprises, using both the *SUE* and *CAR* measures of earnings surprises. Consistent with the portfolio sorts in Table 4, short interest predicts subsequent earnings surprises and its predictive power once again increases as short sale constraints become more binding (low *RIO*). The coefficient on *Short* is consistently negative and significant at the 1% level (except in one case, for *CAR* regressions with mid-level *RIO*) and becomes larger for stocks with low *RIO.* We also note that past return momentum and previous quarter's earnings surprises have significant positive associations with future earnings surprises, a result consistent with Chan, Jegadeesh, and Lakonishok (1996). Using standardized decile ranks (Table A2 in the Appendix), the coefficient of short interest in the CAR equation is -0.67 (t-value = -4.93). This implies that there is a 0.67% return difference around earnings announcements between stocks in the highest and lowest short interest deciles.

Overall, the results obtained from the Fama-MacBeth regressions in Table 5 confirm that the level of short interest is informative about future earnings news in the cross-section of US stocks, especially when short selling is more constrained. Our results also suggest that short interest contains information about future earnings surprises, even after controlling for the previous quarter's surprise and stock return momentum.

3.4. Short interest and future changes in analyst forecasts

Changes in analysts' earnings forecasts are often regarded as a proxy for changes in firm fundamentals. If short sellers are able to predict deteriorating fundamentals before analysts' actions affect prices, we expect to find a significant relationship between short interest and future changes in consensus earnings-per-share (*EPS*) forecasts, especially when shorting constraints are more binding. We test whether the level of short interest during the previous month, *t-1*, can predict the change in the mean *EPS* forecast between the end of month *t-1* and the end of month *t*. We examine the relation between current levels of short interest and two measures of future forecast changes. The first measure *($\Delta Forecast1$)* is the one-month-ahead change in the consensus *EPS* forecast, normalized by the absolute value of mean *EPS* forecast at the end of month *t-1* and reported in percentage form: $\Delta EPS_t/|EPS_{t-1}|$. The second measure *($\Delta Forecast2$)* is the same change in *EPS*, normalized by the share price at the end of month *t-1*: $\Delta EPS_t/P_{t-1}$.

Table 6 presents the results obtained using the calendar time portfolio method. The analysis in this section is based on the sub-sample of stocks with at least one analyst forecast reported by First Call. In Panel A, stocks are sorted into quintiles based upon the level of short interest at month *t-1.* The first column in Panel A shows the mean level of short interest within each quintile. The next two columns verify that the predictability of stock returns is preserved in this particular sub-sample. The last two columns show that the level of short interest predicts future changes in analysts' consensus earnings forecasts in the cross-section. Stocks in the highest short interest quintile are followed by a more negative change in analyst forecast when compared to those in the lowest quintile and this difference is statistically significant for both measures of earning forecast changes. For example, using the *ΔForecast1* measure, which can be interpreted as percentage change in projected earnings, we find a difference of -5.31% (t-value = -4.95) between stocks in the top short interest quintile and those in the bottom quintile. Thus, short sellers appear to take positions on firms that are about to be revised downward by analysts based on their expectations of firms' future earnings performance. We show in Panel B that the predictive power of short interest increases as shorting costs become significantly more severe.

Fama-MacBeth regressions, presented in Table 7, provide a similar picture. The coefficient on *Short* is negative and significant in all panels and specifications. We obtain similar results when independent variables are transformed into decile ranks and then standardized (see Table A3 in the Appendix). For example, controlling for same firm characteristics as in Table 7, the difference in *ΔForecast1* (percentage change in consensus EPS estimate) between the highest and lowest short interest decile is -4.3% (t-value = -6.62). Overall, the results in Tables 6 and 7 suggest that short selling predicts future analyst forecast revisions.

3.5. Longer-term results

If short sellers are trading based on their superior ability to identify changes in fundamentals, the time pattern of their holdings should reflect this informational advantage. Specifically, we expect short interest to increase ahead of bad news and decrease ahead of good news. To examine these time trading dynamics, we compare the size of short sellers' relative positions around news events starting 12 months before and ending 12 months after each event. The relative positions are computed as the difference in short interest between "good-news" stocks and "bad-news" stocks. We define "good-news" and "bad-news" stocks based upon the subsequent occurrence of an event that alters the market's perception of the stock's fundamental value.

Figure 1 depicts the inter-temporal evolution of short sellers' relative position surrounding earnings surprises and analyst forecast revisions. We use the same two definitions of earnings surprises (*SUE*, *CAR*) and forecast revisions (*ΔForecast1, ΔForecast2*) as before. We first group stocks into quintiles based on these four variables. Sorting is done cross-sectionally each quarter (for earnings surprises) and each month (for earnings forecast revisions). We then calculate the difference between short interest for stocks in the bottom quintile (bad news) and those in the top quintile (good news). Consistently, across the four measures we find that bad-news stocks gradually build up much higher short positions than good-news stocks. This process evolves gradually over the year before the news event and reverses over the year following the event.

In Figure 2, we repeat the same analysis using the two public news measures from RavenPack, *NEWS1* and *NEWS2*. Due to data availability, the sample period covered here is shorter, from 2000 to 2010. However, the observed short interest pattern is very similar to that in Figure 1. Once again, short sellers gradually increase their positions in stocks that are about to experience bad news and then unwind these positions after the news become public.

We next examine the "shelf life" or duration of the informational content of short interest. We tabulate returns and news for future monthly holding periods, up to one year into the future. The results are presented in Table 8. The predictive power of short interest for returns increases each month for the year following the sorting period. For example, the difference in news content between highly shorted and least shorted stocks (using the NEWS1 measure) appears to be monotonically increasing with the event horizon. At the 12-month horizon, this difference in news content is -0.736 (t=-7.4). We observe a similar monotonic pattern for stock returns, earnings, and analyst forecast revisions.

Consistent with the graphical analysis in Figures 1 and 2, the results in Table 8 show that short sellers begin to adjust their positions up to one year before news events. Thus, short sellers' knowledge is informative for returns and news occurring up to one year into the future. These results highlight the importance of looking at periods longer than a few trading days when assessing the informational content of short sale transactions.

4. Relation between return predictability and fundamental value predictability

Thus far in the paper we have established that short interest predicts future changes in firm fundamental value. This is a novel result in the literature, which complements the more established result that short interest predicts future stock returns. In this section we seek to determine if these two results are related. Specifically, we hypothesize that short interest's return predictability established in prior research is caused, at least in part, by the informational content of short sale positions. We use four different approaches to address this question empirically.

4.1. Fama-MacBeth regressions

We begin with estimating Fama-MacBeth regressions that are similar to those reported in Tables 3, 5, and 7, with the following two exceptions: (i) the dependent variable is now excess returns, and (ii) contemporaneous changes in firm fundamentals are added as independent variables, alongside lagged short interest and controls for size, book-to-market, and 12-month return momentum. Our focus is on measuring short interest's explanatory power for returns, after controlling for short sellers' information about news events. We run regressions with and without firm fundamental variables and compare the coefficient on short interest between the two specifications. If the coefficient on short interest loses importance when fundamental variables are included, it means that short interest's ability to predict returns is explained (at least in part) by information about these firm fundamentals. The results are presented in Table 9.

Model 1 in Table 9 provides the baseline regression model, which does not include fundamental variables. The negative coefficient of short interest in this baseline regression (-0.049, t=-2.80) confirms that short interest significantly predicts the cross-section of stock returns in this sample. When we add the three fundamental variables that short sellers can apparently predict (analyst forecasts revisions, earnings surprises, and news sentiment), the coefficient on short interest declines (in magnitude) to -0.035 (t=-1.86) in Model 2, and to -0.025 (t=-1.48) in Model 3. In both models (2 and 3), the coefficient is substantially smaller in magnitude than the baseline coefficient, and is no longer significant. The difference between coefficients on short interest in Model 1 and 2 (3) is highly significant with a t-value of 4.67 (3.08). Comparing the magnitude of these coefficients, the results suggest that approximately half of short sellers' ability to predict returns can be explained by their ability to predict future fundamentals. We obtain similar results in regressions based on decile ranks (reported in Table A4 of the Appendix).

4.2. Forthcoming news and return predictability

In our second approach, we use a more direct approach to determine if short sellers' knowledge of forthcoming news is related to short interests' return predictability. If short sellers have no information about future news, their ability to predict returns should not be related to news flow. But if short sellers are primarily informed about future news, the relation between short interest and returns should be stronger when news are forthcoming, particularly if these news are negative.

The test consists of regressing excess returns at time t on past *short interest* (measured at t-1) and current *news content* (measured at time t), accounting for the usual risk control variables (size, book-to-market, and return momentum). We construct the following three variables to measure the *news content* at time *t* :

- *News_Dummy* is a news indicator that is equal to 1 if any news event occurs during month (*t*), following the month (*t-1*) when short interest is observed.
- *Neg_NEWS1_Dummy* (and, alternatively, *Neg_ NEWS2_Dummy*) are indicator variables equal to 1 only if the respective values of *NEWS1* or *NEWS2* are negative (i.e., the news measures are below median).
- *Pos_NEWS1_Dummy* (and, alternatively, *Pos_NEWS2_Dummy*) are indicator variables equal to 1 only if the respective values of *NEWS1* or *NEWS2* are positive (i.e., the news measures are above median).

The results are presented in Table 10. The first (leftmost) column includes the baseline regression without *news* dummies. The negative coefficient on *short* confirms the well-known predictive relation between short interest and future returns. The coefficient of short interest in the baseline regression is -0.045 (*t=-2.26*). The next two columns include the *news* dummy variables, along with interactions between *news* and *short*. The second (centermost) column uses the unsigned *news* dummy variable, while the third (rightmost) column uses the two signed *news* dummy variables (positive and negative). We hypothesize that the presence of news, and particularly bad news, increases the predictive power of short interest for future returns. Thus, we expect to find a negative coefficient on the interactions between *short* and *news* dummies.

In the second column of Table 10, the coefficient on *short* drops (in magnitude) from -0.045 to -0.027, and is no longer significant (t=-1.27). By contrast, the coefficient on the interaction term (*short*news_dummy*) is negative, large in magnitude, and significant (-0.050, t=-3.98). Taken together, these results imply that the relation between short interest and future returns is insignificant in the absence of forthcoming news, the coefficient of *short* being -0.027. However, in the presence of news, the sum of the coefficient on *Short* and the interaction term (-0.077) represent the total effect of short interest on returns. This effect is large in magnitude and statistically significant.

The third column of Table 10 documents an asymmetry between positive and negative news. The predictive relation between short interest and stock returns appears to be driven exclusively by *short*'s ability to predict bad news, as evidenced by the negative and highly significant coefficient on the term interacting short interest with the bad news dummy (-0.092, t=-4.74).

Regressions using standardized decile ranks provide some economic intuition about the magnitudes of the results (see Table A5 in the Appendix). In the baseline model without fundamental variables, the return difference between highest and lowest short interest decile is -82.7 basis points per month and significant at 5% level. In the news model, this same difference diverges to -55.8 basis points in the absence of forthcoming news, and to -126.3 basis points when news are forthcoming. The spread is even larger when forthcoming news are negative. Thus, the predictive power of short interest for future returns more than doubles in the presence of news.[9]

4.3. Explanatory power of short interest

In our third approach, we measure short interest's contribution to the explanatory power of the return regressions in Table 9. Specifically, we compute the percentage difference in average adjusted R^2 between models with, and without short interest (not

---

[9] We repeat the analysis in Table 10 using only analyst forecast changes as the news item. The results (unreported) are very similar and are available upon request. We report the results with general news items since they cover all news categories.

tabulated). We find that adding short interest to the model increases adjusted R^2 by 11.16% in Model 1 of Table 9. This implies that short interest explains an incremental 11.16% of the variation in stock returns when none of the fundamental news variables is included. In contrast, the contribution of short interest is only 5.25% and 1.58% in Models 2 and 3, respectively, when fundamental variables are also present in the regressions. The differences are significant at the 1% level. This means that the predictive power of short interest for future stock returns is strongly related to firms' fundamental performance. After controlling for fundamental performance, the informational content of short interest for future stock returns is small because most of it is subsumed by future changes in fundamental performance.

4.4. Time series test

For our fourth and final approach, we conduct a time-series test to determine if the ability of short interest to predict future returns is related to its ability to predict fundamental information (results are not tabulated). We estimate two sets of monthly Fama-MacBeth regressions: in the first one short interest in month *t-1* is regressed on excess stock return in month *t*, calculated as the difference between raw return and one-month T-bill rate. In the second regression short interest at *t-1* is regressed on future news releases (*NEWS2*), earnings surprises (*CAR*), and analyst forecast revisions (*ΔForecast2*). We then run a time-series regression between the monthly R^2s obtained from the above two sets of cross-sectional regressions using the R^2s from the return regressions as the dependent variable. The result from the time-series regression is the following, with t-values in parenthesis:

$$R^2 \textit{ (return regressions)} = -0.0103\times \textit{Intercept} + 0.3397\times R^2 \textit{ (fundamental regressions)}$$
$$(t=-7.15) \qquad\qquad (t=2.79)$$

The R^2 values from the return regression are significantly positively related to the R^2 values from the fundamental regressions. During months when short interest better predicts fundamental news, it also predicts returns better.

5. Robustness checks

We conduct four additional tests to assess the robustness of our results.

5.1. Varying filters for NEWS1 and NEWS2.

Our main public news definitions, NEWS1 and NEWS2, rely on RavenPack's sentiment scores based on computerized content analysis. Although we show in Panel B of Table 2 that – from a market perspective – these news are unexpected and correctly signed, we repeat our analysis in Table 3 using more restrictive filters for defining NEWS1 and NEWS2. Specifically, we require a minimum absolute return of 0.5 % (and alternatively 1.0%) in excess of market return before we categorize a day as "bad news" or "good news." With this new approach, news days associated with abnormal stock returns whose magnitude is lower than the new filter (0.5% or 1%) are categorized as neutral. The results are shown in Table A6 of the Appendix. Short interest continues to predict news and the magnitudes of the *short* coefficients remain similar to those reported in Table 3. Our results are robust to this alternative classification of news.

5.2. Persistence of public news

If news are persistent over time, our results could be driven by short sellers' ability to better interpret past news and take positions accordingly. Under this scenario, short interest would simply reflect the information content of news that are already public. To determine if our results are driven by news persistence we repeat our analysis in Table 3, controlling for past news up to six months earlier. The results (untabulated) show that short interest continues to predict future news even after controlling for past news over much longer time periods.

5.3. Alternative news source

To assure that our *news* results are not specific to the RavenPack dataset, we use Chan's (2003) news data that covers a random sample of approximately one-quarter of CRSP stocks during the period from 1980 to 2000. In Chan's data, news items are collected from the Dow Jones Interactive Publications Library. We compute market excess return for each news day and use only news with returns larger than 1% in absolute value. As with RavenPack news, we find that short interest continues to predict future news and that results are stronger for stocks that are more costly to short. As in previous tests, we continue to find that the predictive power of short interest significantly diminishes once we control for forthcoming news. Overall, the use of Chan's news database demonstrates the robustness of our results to alternative news databases.

5.4. Different proxies for short sale constraints

In the main results we use *residual institutional ownership* (RIO) as a proxy for shorting costs. We repeat our analysis using four alternative proxies: market capitalization (SIZE), Amihud's (2002) measure of illiquidity, share turnover, and analyst coverage. The results (untabulated) show that the relation between short interest and future news is stronger among stocks that are small, more illiquid, lightly traded, or followed by fewer analysts. Overall, these results corroborate our earlier conclusion that short interest's ability to predict future news is stronger in the presence of higher shorting costs, consistent with the predictions of Diamond and Verrecchia (1987).

6. Conclusion

Short interest predicts stock returns in the cross-section because short sellers are informed traders who generate value-relevant information. We show that short sellers correctly anticipate negative earnings surprises, bad public news, and downgrades in analyst earnings forecasts several months ahead. Their ability to predict future fundamental events appears to be the dominant driver of their ability to predict future returns. Shorts seem to be particularly well informed about stocks with low levels of institutional ownership, which are presumably harder to short.

Our results suggest that short sellers' advantage over other traders comes from their better ability to predict future material events that affect firm value, rather than shorter-horizon trading strategies. This finding has important policy implications, because it raises questions about the source of this information about future events. Possibly short sellers run the same prediction models that analysts and news reporters do, but are just quicker or more effective in trading on this information compared to others. Alternatively, short sellers might trade on rumors or information leaks to which other traders do not have access. Elucidating the source of short sellers' informational advantage remains an important issue for future research.

References

Amihud, Yakov, 2002, Illiquidity and stock returns: cross-section and time-series effects, Journal of Financial Markets 5, 31-56.

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

Beatty, R., 1989. Auditor reputation and the pricing of initial public offerings. The Accounting Review 64, 693-709.

Beatty, R., Ritter, J., 1986. Investment banking, reputation, and the underpricing of initial public offerings. Journal of Financial Economics 15, 213-232.

Billett, M., Flannery, M., Garfinkel, J., 1995. The effect of lender identity on a borrowing firm's equity return. Journal of Finance 50, 699-718.

Boehmer, E., Kelley, E., 2009. Institutional investors and the informational efficiency of prices. Review of Financial Studies 22, 3563-3594.

Boehmer, E., Jones, C., Zhang, X., 2008. Which shorts are informed? Journal of Finance 63, 491-527.

Boehmer, E., Jones, C., Zhang, X., 2013. Shackling short sellers: The 2008 shorting ban. Review of Financial Studies, forthcoming.

Boehmer, E., Wu, J., 2013. Short selling and the price-discovery process. Review of Financial Studies 26, 287-322.

Boehmer, E., Huszar, Z., Jordan, B., 2010. The good news in short interest. Journal of Financial Economics 96, 80-97.

Brennan, M., Subramanyam, A., 1995. Investment analysis and price formation in securities markets. Journal of Financial Economics 38, 361-381.

Brown, S., Warner, J., 1985. Using daily stock returns: the case of event studies. Journal of Financial Economics 14, 3-31.

Carhart, M., 1997. On persistence in mutual fund performance. Journal of Finance 52, 57-82.

Carter, R., Manaster, S., 1990. Initial public offerings and underwriter reputation. Journal of Finance 45, 1045-1067.

Chan, L., Jegadeesh, N., Lakonishok, J., 1996. Momentum strategies. Journal of Finance 51, 1681-1713.

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

Christophe, S., Ferri, M., Hsieh, J., 2010. Informed trading before analyst downgrades: Evidence from short sellers. Journal of Financial Economics 95, 85-106.

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

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

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

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

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

Diether, K., Lee, K., Werner, I., 2009. Short-sale strategies and return predictability. Review of Financial Studies 22, 575-607.

Engelberg, J., Reed, A., Ringgenberg, M., 2012. How are shorts informed?: Short sellers, news, and information processing. Journal of Financial Economics 105, 260-278.

Fama, E., French, K., 2002. Testing trade-off and pecking order predictions about dividends and debt. Review of Financial Studies 15, 1-33.

Fama, E., French, K., 1992. The cross-section of expected stock returns. Journal of Finance 47, 427-465.

Fama, E., MacBeth, J., 1973. Risk, return and equilibrium: Empirical tests, 1973. Journal of Political Economy 81, 607-636.

Foster, G., Olsen, C., Shevlin, T., 1984. Earnings releases, anomalies, and the behavior of security returns. The Accounting Review 59, 574-603.

Henry, T., Koski, J., 2010. Short selling around seasoned equity offerings. Review of Financial Studies 23, 4389-4418.

Hirshleifer, D., Teoh, S., Yu, J., 2011. Short arbitrage, return asymmetry and the accrual anomaly. Review of Financial Studies 24, 2429-2461.

James, C., 1987. Some evidence on the uniqueness of bank loans. Journal of Financial Economics 19, 217-236.

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

Karpoff, J., Lou, X., 2010. Short sellers and financial misconduct. Journal of Finance 65, 1879-1913.

Lummer, S., McConnell, J., 1989. Further evidence on the bank lending process and the capital-market response to bank loan agreements. Journal of Financial Economics 25, 99-122.

Mikkelson, W., Partch, M., 1986. Valuation effects of security offerings and the issuance process. Journal of Financial Economics 15, 31-60.

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

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

Saffi, P., Sigurdsson, K., 2011. Price efficiency and short selling. Review of Financial Studies 24, 821-852.

Solomon, D., 2012. Selective publicity and stock prices. Journal of Finance 67, 599-638.

Table 1. Short interest and future returns

The sample contains common stocks listed on the NYSE, Amex, and Nasdaq during the period from January 1988 to December 2010. Each month, we sort stocks into short interest deciles based on $Short_{t-1}$, the ratio of short interest to shares outstanding. This table shows, for each decile, the time-series means of short interest ($Short_{t-1}$), holding-period raw returns (Ret_t), and abnormal monthly returns (α_t) computed as the intercept from a four-factor model including *mktrf*, *smb*, *hml*, and *umd*. All returns are measured over a one month horizon, during the month following portfolio formation. For each group, we show raw returns, abnormal returns, and level of short interest. T-statistics are shown in italics and are based on robust standard errors. *, **, and *** denote significance level at the 10%, 5%, and 1% level, respectively.

	$\mathbf{Short_{t-1}}$	$\mathbf{Ret_t}$	$\boldsymbol{\alpha_t}$
Short1 (low)	0.009	1.19***	0.38***
	0.115	1.33***	0.37***
	0.375	1.32***	0.28***
	0.713	1.27***	0.19**
	1.114	1.16***	0.07
	1.614	0.99**	0.09
	2.241	1.07**	-0.02
	3.162	0.97*	-0.09
	4.893	0.69	-0.35***
Short10 (high)	11.788	0.46	-0.57***
Short10 – Short1	11.779***	-0.73***	-0.95***
t-value	*18.89*	*-2.05*	*- 4.36*

Table 2. Short interest as predictor of public news: a portfolio approach

The sample contains common stocks listed on the NYSE, Amex, and Nasdaq during the period from January 2000 to December 2010. We also require that each stock be included in RavenPack news database. We report the time series averages of news content and return in month t based on short interest in month t-1. We infer news content from the sentiment score provided by Ravenpack. *NEWS1* is the average content of news measure where news are classified as either negative, neutral, or positive. Then we assign numerical values of +1 to all positive news items, zero to all neutral items, and –1 to all negative news items and compute a monthly aggregate news measure by averaging the (signed) values of all news items that month. *NEWS2* is the average of sentiment score of news in a month. Months with no news items are assigned an aggregate *NEWS1* (*NEWS2*) measure equal to 0 (50). All news measures are orthogonalized with respect to firm size. Abnormal monthly returns (α) are computed as the intercept from a four-factor model including *mktrf*, *smb*, *hml*, and *umd*. All returns are measured over a one month horizon, during the month following portfolio formation. In Panel A we present time series averages of cross sectional descriptive statistics. In Panel B, we present time series averages of daily mean and median values of percentage daily excess returns for no news, positive, negative and neutral news stocks. In Panel C, we conduct a univariate sort on short interest ($Short_{t-1}$). In Panel D we conduct an independent double sort, first on residual institutional ownership (RIO_{t-1}) and then on short interest ($Short_{t-1}$). Ret_t is monthly raw return; RIO_{t-1} is the residual from a quarterly regression of institutional ownership, *IO*, on firm size (*Size*). $Size_{t-1}$ is the market value of equity and defined as share price times the number of shares outstanding. Institutional ownership (IO_{t-1}) is the percentage of shares owned by institutions as reported in 13F filings. T-statistics are shown in italics and are based on robust standard errors. *, **, and *** denote significance level at the 10%, 5%, and 1% level, respectively.

Panel A: Descriptive Statistics

	Mean	**Median**	**Std**	**P10**	**P90**
$SIZE_{t-1}$	4.92	0.76	19.65	0.11	8.99
BM_{t-1}	0.53	0.41	0.59	0.12	1.03
R_MOM_{t-1} (%)	26.94	9.47	81.8	-32.8	93.7
IO_{t-1} (%)	65.92	70.2	26.4	27.08	95.61
$NEWS1_t$	0	-0.07	0.43	-0.37	0.64
$NEWS2_t$	0	-1.64	9.21	-7.02	10.9
RET_t(%)	0.86	0.41	13.14	-13.5	15.46
$SHORT_{t-1}$	4.88	3.3	5.53	0.42	11.05

Panel B: Average Daily Excess Returns of No News, Neutral, Bad and Good News Stocks

		News 1			News 2		
	No News	Bad	Neutral	Good	Bad	Neutral	Good
Mean	0.069	-1.418	0.028	1.095	-1.423	-0.004	1.318
Median	-0.041	-0.947	-0.017	0.529	-0.965	-0.013	0.701

Panel C: Future news count as a function of current short interest

	$Short_{t-1}$	$NEWS1_t$	$NEWS2_t$	Ret_t	α_t
Short1 (low)	0.15	0.031***	0.693***	1.5***	0.85***
	0.75	0.022***	0.317***	1.2***	0.49***
	1.39	0.011***	0.126***	1.13**	0.42***
	2.07	0.013***	0.183	1.04**	0.31**
	2.87	0.007**	0.113	1.03*	0.29**
	3.75	0.001	0.05	0.92*	0.19
	4.82	-0.068**	-0128**	0.71	0.03
	6.37	-0.015***	-0.237***	0.5	-0.11
	8.98	-0.029***	-0.541***	0.39	-0.22
Short10 (high)	17.6	-0.033***	-0.673***	0.19	-0.54***
Short10- Short1	17.45	-0.064	-1.366	-1.31	-1.39
t-value	*22.8*	*-8.38*	*-8.05*	*-2.98*	*4.72*

Panel D: Future news count as a function of short interest and residual institutional ownership

***News* Definition**		**$RIO1_{t-1}$ (low)**	**RIO2t-1 (medium)**	**$RIO3_{t-1}$ (high)**	**RIO1-RIO3**	**t-value**
	$Short1_{t-1}$	0.027***	0.027***	0.019***	0.008	*1.24*
		-0.002	0.016***	0.017***	-0.15	*-2.08*
		-0.011**	0.007**	0.015***	-0.026	*-2.79*
***NEWS1*$_t$**		-0.029***	-0.003	-0.005	-0.024	*-2.54*
	$Short5_{t-1}$	-0.059***	-0.035***	-0.014***	-0.045	*-4.02*
	Short1-Short5	-0.086	-0.062	-0.033	-0.53	
	t-value	*-8.03*	*-7.38*	*-4.22*	*-5.89*	
	$Short1_{t-1}$	0.56***	0.51***	0.263**	0.297	*2.12*
		0.0	0.24***	0.187**	-0.187	*-1.46*
		-0.169*	0.153***	0.261***	-0.43	*-2.27*
***NEWS2*$_t$**		-0.52***	-0.038	-0.062	-0.45	*-2.36*
	$Short5_{t-1}$	-1.1***	-0.63***	-0.21***	-0.89	*-3.58*
	Short1-Short5	-1.66	-1.14	-0.476	-1.184	
	t-value	*-7.09*	*-6.68*	*-3.25*	*-5.57*	

Table 3. Short interest as predictor of public news: A regression approach

The sample contains common stocks listed on the NYSE, Amex, and Nasdaq during the period from January 2000 to December 2010. We also require that each stock be included in RavenPack news database. The table shows monthly Fama-MacBeth regressions where the dependent variable is content of future monthly news items (*NEWS*). We report the time series averages of month t news content and return based on short interest in month *t-1*. We infer news content from the sentiment score provided by Ravenpack. *NEWS1* is the average content of news measure where news are classified as either negative, neutral, or positive. Then we assign numerical values of +1 to all positive news items, zero to all neutral items, and −1 to all negative news items and compute monthly aggregate news measure by averaging the (signed) values of all news items that month. *NEWS2* is the average of sentiment score of news in a month. Months with no news items are assigned an aggregate *NEWS1* (*NEWS2*) measure equal to 0 (50). All news measures are orthogonalized with respect to firm size. The independent variables are *Short* (the ratio of short interest to shares outstanding), *RIO* (residual institutional ownership , where institutional ownership is the percentage of shares owned by institutions as reported in 13F filings and residuals are obtained from quarterly regressions of institutional ownership on size), *Size* (the natural logarithm of the market value of equity), *BM* (the natural logarithm of the ratio between book value of equity and the market value of equity calculated at least three months before the short interest data), *R_MOM* (the mean stock return performance over the previous 12 months) and $News_{t-1}$ (*news content* measured in month t-1). The table reports time-series averages of the coefficient estimates. T-statistics are shown in italics below each coefficient and are based on robust standard errors.

	All	Low RIO	Med RIO	High RIO	All	Low RIO	Med RIO	High RIO
	NEWS1$_t$	***NEWS1***$_t$	***NEWS1***$_t$	***NEWS1***$_t$	***NEWS2***$_t$	***NEWS2***$_t$	***NEWS2***$_t$	***NEWS2***$_t$
Intercept$_{t-1}$	0.002	0.024	0.022	-0.009	-0.067	0.512	0.497	-0.638
	0.26	*1.66*	*1.33*	*-0.37*	*-0.33*	*1.74*	*1.47*	*-1.28*
Short$_{t-1}$	-0.309	-0.552	-0.416	-0.177	-5.633	-10.854	-7.485	-2.861
	-9.38	*-10.18*	*-7.15*	*-6.46*	*-8.10*	*-9.06*	*-6.96*	*-4.65*
Size$_{t-1}$	0.000	0.000	0.000	-0.001	-0.004	-0.023	-0.030	0.003
	0.44	*-0.32*	*-0.35*	*-0.43*	*-0.31*	*-1.02*	*-1.54*	*0.09*
BM$_{t-1}$	-0.010	-0.005	-0.010	-0.017	-0.311	-0.199	-0.299	-0.464
	-2.68	*-1.17*	*-2.69*	*-5.11*	*-3.76*	*-2.07*	*-3.74*	*-6.44*
R_MOM$_{t-1}$	0.051	0.040	0.060	0.070	1.448	1.162	1.711	1.966
	8.67	*7.62*	*7.30*	*8.83*	*8.87*	*8.54*	*7.79*	*8.62*
RIO$_{t-1}$	0.048	0.038	0.026	0.058	1.025	0.802	0.350	1.119
	6.50	*3.16*	*0.80*	*4.09*	*7.23*	*3.11*	*0.51*	*4.01*
News$_{t-1}$	0.002	0.000	0.002	0.001	-0.005	0.001	-0.008	-0.012
	0.67	*0.08*	*0.52*	*0.22*	*-1.36*	*0.11*	*-1.76*	*-2.77*
Adj. R^2	1.31%	1.36%	1.29%	1.41%	1.72%	1.73%	1.74%	2.04%

Table 4. Short interest as predictor of earnings surprises: A portfolio approach

The sample contains common stocks listed on the NYSE, Amex, and Nasdaq during the period from January 1988 to December 2010. In Panel A, we sort stocks into quintiles based on standardized unexpected earnings, (*SUE*), defined as the change in earnings per share from quarter *q*-4 to quarter *q* divided by the standard deviation of unexpected earnings over the last eight quarters. We report time-series means of SUE_q and announcement abnormal returns, CAR_q, defined as the average (across firms) of the daily market-adjusted cumulative abnormal return (in percent) during the [-2, +2] window around the earnings announcement date. In Panel B, we sort on quarter *q-1* short interest ($Short_{q-1}$) and report the time-series average of one quarter ahead earnings surprises, SUE_q and CAR_q. *Short* is the ratio of short interest to shares outstanding. In Panels C and D we conduct an independent double sort, first on quarter *q-1* residual institutional ownership (*RIO*) and then on quarter *q-1* short interest (*Short*). Panel C reports the corresponding CAR_q for the next-quarter earnings announcement, and Panel D reports the average SUE_q. RIO_{q-1} is the residual from a quarterly regression of institutional ownership, *IO* on firm size (*Size*). *Size* is the market value of equity and defined as share price times the number of shares outstanding. Institutional ownership (*IO*) is the percentage of shares owned by institutions as reported in 13F filings. T-statistics are shown in italics and are based on robust standard errors. *, **, and *** denote significance level at the 10%, 5%, and 1% level, respectively.

Panel A: Descriptive statistics on earnings surprises

	$SUE1_q$ (low)				**$SUE5_q$ (high)**	**SUE5 – SUE1**	**t-value**
SUE_q	-1.98	-0.22	0.13	0.57	2.11	4.09***	*36.22*
CAR_q	-1.65	-0.95	0.25	1.05	1.55	3.20***	*23.21*

Panel B: Future earnings surprises as a function of current short interest

	$Short_{q-1}$	**SUE_q**	**CAR_q**
Short 1 (low)	0.086	0.115	0.310
	0.651	0.186	0.095
	1.545	0.192	0.007
	2.903	0.132	0.033
Short 5 (high)	8.556	0.020	-0.126
Short 5 – Short 1	8.470***	-0.095***	-0.436***
t-value	*9.50*	*-3.19*	*-4.07*

Panel C: Future earnings-event CARs a function of current short interest and residual institutional ownership

	RIO_{q-1} 1 (low)	**RIO_{q-1} 2 (medium)**	**RIO_{q-1} 3 (high)**	**RIO1 – RIO3**	**t-value**
***$Short_{q-1}$* 1**	0.11	0.49	0.37	-0.26**	*-2.27*
	-0.14	0.16	0.25	-0.39***	*-2.85*
	-0.20	0.04	0.13	-0.33**	*-2.33*
	-0.28	0.09	0.23	-0.51***	*-3.71*
***$Short_{q-1}$* 5**	-0.58	-0.03	0.03	-0.61***	*-3.22*
Short1 - Short5	0.69***	0.52***	0.34***		
t-value	*3.95*	*3.21*	*2.77*		

Panel D: Future SUE as a function of current short interest and residual institutional ownership

	RIO_{q-1} 1 (low)	**RIO_{q-1} 2 (medium)**	**RIO_{q-1} 3 (high)**	**RIO1 – RIO3**	**t-value**
$Short_{q-1}$ 1	0.17	0.08	0.04	0.13***	*3.53*
	0.25	0.18	0.12	0.13***	*4.36*
	0.20	0.21	0.15	0.05	*1.65*
	0.12	0.17	0.11	0.01	*0.07*
$Short_{q-1}$ 5	-0.11	0.04	0.07	-0.18***	*-4.57*
1 – 5	0.28***	0.04	-0.03*		
t-value	*7.46*	*1.05*	*-1.76*		

Table 5. Short interest as predictor of earnings surprises: A regression approach

The sample contains common stocks listed on the NYSE, Amex, and Nasdaq during the period from January 1988 to December 2010. We estimate quarterly Fama-MacBeth regressions using earnings surprises as the dependent variable. We split the data every quarter based on residual institutional ownership (*RIO*) and run separate regressions for the top 30%, middle 40%, and bottom 30% of the data. Earnings surprises are measured using either *SUE* or *CAR*. SUE_q is the change in quarterly earnings per share from quarter q-4 to quarter q, divided by the standard deviation of unexpected earnings over the last eight quarters. CAR_q is the daily market-adjusted cumulative abnormal return (in percent) during the [-2, +2] window around the earnings announcement date. The independent variables are $Short_{q-1}$ (the ratio of short interest to shares outstanding), RIO_{q-1} (the residual institutional ownership as defined in Table 3), $Size_{q-1}$ (the natural logarithm of the market value of equity), BM_{q-1} (the natural logarithm of the ratio between book value of equity and the market value of equity calculated at least three months before the short interest data), and R_MOM_{q-1} (the mean stock return performance over the previous 12 months), E_MOM_{q-1} (previous quarter's earnings surprise). The table reports time-series averages of the coefficient estimates. T-statistics are shown in italics below each coefficient and are based on robust standard errors.

	All	Low RIO_{q-1}	Med RIO_{q-1}	High RIO_{q-1}	All	Low RIO_{q-1}	Med RIO_{q-1}	High RIO_{q-1}
	SUE_q	SUE_q	SUE_q	SUE_q	CAR_q	CAR_q	CAR_q	CAR_q
$Intercept_{q-1}$	-0.592	-0.428	-0.509	-0.871	0.229	0.189	0.481	0.304
	-9.83	*-4.37*	*7.63*	*-9.19*	*1.01*	*0.62*	*1.84*	*0.71*
$Short_{q-1}$	-0.016	-0.031	-0.016	-0.010	-0.045	-0.079	-0.028	-0.032
	-5.87	*-6.50*	*-4.59*	*-3.40*	*-4.12*	*-4.95*	*-1.14*	*-2.97*
$Size_{q-1}$	0.045	0.035	0.039	0.061	-0.004	0.016	-0.018	-0.014
	7.48	*4.20*	*5.79*	*8.85*	*-0.21*	*0.70*	*-0.84*	*-0.38*
BM_{q-1}	-0.105	-0.110	-0.109	-0.099	0.152	0.225	0.192	0.005
	-6.29	*-5.04*	*-6.20*	*-5.08*	*3.20*	*4.06*	*2.85*	*0.08*
R_MOM_{q-1}	0.004	0.003	0.004	0.005	0.000	0.001	-0.001	0.001
	9.18	*6.65*	*9.19*	*10.94*	*0.37*	*1.10*	*-0.51*	*0.67*
RIO_{q-1}	0.001	0.000	-0.000	0.002	0.008	0.014	0.018	0.001
	1.11	*0.37*	*-0.44*	*3.36*	*4.37*	*3.89*	*3.70*	*0.20*
E_MOM_{q-1}	0.300	0.304	0.303	0.285	0.008	0.009	0.006	0.011
	25.33	*21.67*	*20.95*	*21.83*	*1.55*	*1.56*	*0.83*	*1.93*
Adj. R^2	13.91%	13.20%	14.466%	14.75%	0.84%	1.47%	0.98%	1.08%

Table 6. Short interest as predictor of changes in analysts' earnings forecasts: A portfolio approach

The sample contains common stocks listed on the NYSE, Amex, and Nasdaq that are available in the First Call database with at least one analyst following during the period from January 1990 to December 2010. *Short* is the ratio of short interest to shares outstanding in percentage terms. *Ret* is the one-month ahead raw return, and *IO* is the percentage of shares owned by institutions as reported in 13F filings. ΔEPS_t is the one-month-ahead change in the consensus earnings-per-share (*EPS*) forecast of analysts in cents. *ΔForecast1* is the percentage monthly change in the consensus EPS forecast measured as $\Delta EPSt/|EPS_{t-1}|$. *ΔForecast2* is the monthly change in consensus *EPS* forecast normalized by share price measured as $\Delta EPS_t/P_{t-1}$. In **Panel A**, we sort stocks based on short interest (*Short*), and report future firm performance and change in analysts' earnings per share forecasts (*ΔForecast1 and ΔForecast2*) for the current fiscal year-end. In **Panel B and C,** each month, we sort stocks first on residual institutional ownership (*RIO*) and then on short interest (*Short*). We then report the one-month ahead average change in analysts' earnings per share forecasts (*ΔForecast1 and ΔForecast2*) for the current fiscal year-end. T-statistics are shown in italics and are based on robust standard errors. *, **, and *** denote significance level at the 10%, 5%, and 1% level, respectively.

Panel A: Future performance as a function of current short interest

	$Short_{t-1}$	Ret_t	α_t	*ΔForecast1*	*ΔForecast2*
***Short* 1**	0.40	1.44	0.33**	-4.38	-0.25
	1.22	1.15	0.03	-3.85	-0.21
	2.25	1.13	-0.03	-4.51	-0.20
	3.88	0.98	-0.14	-8.17	-0.44
***Short* 5**	10.17	0.52	-0.62***	-9.69	-0.46
5 – 1	9.77	-0.92	-0.95	-5.31	-0.21
t-value	*19.18*	*-3.50*	*-5.09*	*-4.95*	*-1.98*

Panel B: Future changes in analysts' EPS forecasts as a function of current short interest and residual institutional ownership

		RIO_{t-1} 1 (low)	RIO_{t-1} 2 (medium)	RIO_{t-1} 3 (high)	RIO1 – RIO3	t-value
	***$Short_{t-1}$* 1**	-5.52	-3.40	-7.07	1.55	*1.21*
		-3.28	-2.19	-6.15	2.88***	*3.41*
ΔForecast1		-3.76	-5.16	-5.35	1.59*	*1.85*
		-8.77	-5.89	-9.12	0.35	*0.13*
	***$Short_{t-1}$* 5**	-12.95	-9.66	-7.40	-5.55***	*-3.45*
	Short1 - Short5	7.43***	6.26***	0.33		
	t-value	*4.28*	*4.64*	*0.28*		
	$Short_{t-1}$ 1	-0.29	-0.36	-0.33	0.04	*0.54*
		-0.24	-0.15	-0.22	-0.02	*-0.44*
ΔForecast2		-0.27	-0.24	-0.16	-0.11	*-1.05*
		-0.70	-0.43	-0.25	-0.45	*-1.10*
	$Short_{t-1}$ 5	-0.82	-0.36	-0.32	-0.50***	*-2.96*
	Short1 – Short5	0.53***	0.00	-0.01		
	t-value	*2.87*	*0.04*	*0.22*		

Table 7. Short interest as predictor of changes in analysts' earnings forecasts: A regression approach

The sample contains common stocks listed on the NYSE, Amex, and Nasdaq during the period from January 1990 to December 2010. We estimate monthly Fama-MacBeth regressions using future changes in analysts' earnings forecasts as the dependent variable. We split the data every month based on residual institutional ownership (*RIO*) and run separate regressions for the top 30%, middle 40%, and bottom 30% of the data. Future changes in earnings forecasts are measured as the one-month-ahead change in consensus analyst earnings forecasts. ΔEPS_t is the one-month-ahead change in the consensus EPS forecast in cents. *ΔForecast1* is measured as $\Delta EPS_t/|EPS_{t-1}|$ reported in percentage form. *ΔForecast2* is measured as $\Delta EPS_t/P_{t-1}$.The independent variables are Short (the ratio of short interest to shares outstanding), *RIO* (the percentage of shares owned by institutions as reported in 13F filings), *Size* (the natural logarithm of the market value of equity), *BM* (the natural logarithm of the ratio between book value of equity and the market value of equity calculated at least three months before the short interest data), *E_MOM* (the mean earnings forecast revisions over the previous 12 months), and *R_MOM* (the mean stock return performance over the previous 12 months). Table reports time-series averages of the coefficient estimates. T-statistics are shown in italics below each coefficient and are based on robust standard errors.

	All	Low RIO_{t-1}	Med RIO_{t-1}	High RIO_{t-1}	All	Low RIO_{t-1}	Med RIO_{t-1}	High RIO_{t-1}
	ΔForecast1	***ΔForecast1***	***ΔForecast1***	***ΔForecast1***	***ΔForecast2***	***ΔForecast2***	***ΔForecast2***	***ΔForecast2***
$Intercept_{t-1}$	-41.396	-32.941	-13.802	-50.903	-0.755	-0.380	-1.146	-1.087
	-9.23	*-5.68*	*-0.52*	*-8.93*	*-5.55*	*-2.02*	*0.62*	*-7.80*
$Short_{t-1}$	-0.372	-0.538	-0.383	-0.305	-0.022	-0.033	-0.013	-0.008
	-5.48	*-2.36*	*-2.96*	*-3.31*	*-4.05*	*-2.88*	*-3.38*	*-2.24*
$Size_{t-1}$	2.527	2.022	0.708	3.128	0.045	0.026	-0.081	0.070
	7.61	*5.67*	*0.40*	*7.40*	*5.34*	*2.19*	*-0.65*	*8.71*
BM_{t-1}	-0.889	0.651	-0.314	-2.983	-0.026	-0.023	-0.008	-0.024
	-1.84	*1.16*	*-0.43*	*-3.65*	*-1.97*	*-1.27*	*-0.39*	*-1.05*
R_MOM_{t-1}	0.091	0.085	0.083	0.089	0.002	0.001	0.001	0.001
	8.30	*6.91*	*7.30*	*8.41*	*3.96*	*1.72*	*1.88*	*2.37*
RIO_{t-1}	0.001	-0.039	0.053	0.029	0.002	0.003	0.003	0.001
	0.07	*-1.49*	*0.92*	*0.37*	*2.61*	*1.13*	*1.86*	*0.70*
E_MOM_{t-1}	0.122	0.213	0.203	0.188	0.578	0.683	0.652	0.667
	6.08	*3.20*	*3.97*	*5.67*	*5.11*	*5.27*	*4.30*	*6.54*
Adj. R^2	2.78%	7.25%	4.49%	3.76%	11.15%	15.18%	9.94%	7.47%

Table 8. Short interest as predictor of future fundamental information: A long-term portfolio approach

The sample contains common stocks listed on the NYSE, Amex, and Nasdaq. *NEWS* data covers the period from January 2000 to December 2010 for stocks included in the RavenPack database; whereas analyst forecast and earnings surprise data covers the period from January 1990 to December 2010. Firms are first sorted into deciles based on short interest each month. Sorts are done each quarter for earnings surprise variables. We then show the average difference in percentage return and fundamental performance between highest and lowest short interest portfolios for holding periods up to 12 months. We infer news content from the sentiment score provided by RavenPack. *NEWS1* is the average content of news measure where news are classified as either negative, neutral, or positive .Then we assign numerical values of +1 to all positive news items, zero to all neutral items, and –1 to all negative news items and compute a monthly aggregate news measure by averaging the (signed) values of all news items that month. *NEWS2* is the average of sentiment score of news. Months with no news items are assigned an aggregate *NEWS1* (*NEWS2*) measure equal to 0 (50). All news measures are orthogonalized with respect to firm size. *ΔForecast1* is the percentage change in the consensus analyst EPS forecast. *ΔForecast2* is the change in consensus *EPS* forecast normalized by share price. Earnings surprises are measured using either *SUE* or *CAR*. SUE_q is the change in quarterly earnings per share from quarter q-4 to quarter q, divided by the standard deviation of unexpected earnings over the last eight quarters. CAR_q is the daily market-adjusted cumulative abnormal return (in percent) during the [-2, +2] window around the earnings announcement date. T-statistics are shown in italics below each coefficient and are based on robust standard errors.

	Holding Period in Months											
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Return	-1.313	-2.798	-4.149	-5.492	-6.629	-7.869	-8.954	-9.963	-10.986	-11.967	-12.835	-13.807
t value	*-2.76*	*-3.09*	*-3.15*	*-3.18*	*-3.18*	*-3.22*	*-3.33*	*-3.41*	*-3.37*	*-3.37*	*-3.34*	*-3.34*
NEWS1	-0.065	-0.13	-0.197	-0.26	-0.324	-0.387	-0.446	-0.505	-0.565	-0.621	-0.677	-0.736
t value	*-6.76*	*-6.46*	*-6.64*	*-6.7*	*-6.78*	*-6.71*	*-6.71*	*-6.76*	*-6.84*	*-6.95*	*-7.18*	*-7.4*
NEWS2	-1.265	-2.558	-3.887	-5.087	-6.338	-7.545	-8.67	-9.786	-10.902	-11.935	-12.993	-14.056
t value	*-6.55*	*-6.11*	*-6.22*	*-6.24*	*-6.32*	*-6.25*	*-6.25*	*-6.31*	*-6.37*	*-6.47*	*-6.67*	*-6.89*
ΔForecast 1	-5.35	-8.94	-11.2	-16.12	-21.27	-26.22	-30.57	-34.01	-37.09	-40.15	-42.39	-44.13
t value	*-3.86*	*-4.14*	*-4.23*	*-4.06*	*-4.55*	*-4.62*	*-5.04*	*-5.13*	*-5.28*	*-5.37*	*-5.55*	*-5.33*
ΔForecast 2	-0.219	-0.631	-0.905	-1.2	-1.397	-1.692	-1.867	-2.058	-2.268	-2.381	-2.455	-2.484
t value	*-2.37*	*-2.63*	*-3.15*	*-3.21*	*-3.32*	*-3.15*	*-3.25*	*-3.61*	*-3.84*	*-4.05*	*-4.19*	*-4.25*
SUE			-0.091			-0.157			-0.192			-0.195
t value			*-2.39*			*-2.2*			*-1.87*			*-1.49*
CAR			-0.641			-1.151			-1.503			-1.714
t value			*-3.66*			*-3.33*			*-3.15*			*-3.19*

Table 9. Short interest as predictor of future returns and the role of fundamentals: A regression approach

The sample contains common stocks listed on the NYSE, Amex, and Nasdaq during the period from January 2000 to December 2010. We estimate monthly Fama-MacBeth regressions using excess stock return in month *t*, calculated as the difference between raw return and one-month T-bill rate, as the dependent variable. *ΔForecast1* is the percentage change in the consensus analyst *EPS* forecast in month t. *ΔForecast2* is the monthly change in consensus *EPS* forecast normalized by share price. Earnings surprises are measured using either *SUE* or *CAR*. SUE_t is the change in quarterly earnings per share from quarter *q*-4 to quarter *q*, divided by the standard deviation of unexpected earnings over the last eight quarters. CAR_t is the daily market-adjusted cumulative abnormal return (in percent) during the [-2, +2] window around the earnings announcement date. $NEWS1_t$ is the average content of news measure where news are classified as either negative, neutral, or positive. Then, we assign numerical values of +1 to all positive news items, zero to all neutral items, and –1 to all negative news items and compute a monthly aggregate news measure by averaging the (signed) values of all news items that month. *NEWS2* is the average of sentiment score of news in month t provided by RavenPack. Months with no news items are assigned an aggregate *NEWS1* (*NEWS2*) measure equal to 0 (50). The independent variables are $Short_{t-1}$ (the ratio of short interest to shares outstanding), $Size_{t-1}$ (the natural logarithm of the market value of equity), BM_{t-1} (the natural logarithm of the ratio between book value of equity and the market value of equity calculated at least three months before the short interest data) and R_MOM_{t-1} (the mean stock return performance over the previous 12 months). The contribution of short interest (*CONTR*) is calculated as the percentage difference in average adjusted R^2 between the models with and without short interest in the regressions. The table reports time-series averages of the coefficient estimates. T-statistics are shown in italics below each coefficient and are based on robust standard errors.

	Int_{t-1}	$Short_{t-1}$	$\Delta Forecast1_t$	$\Delta Forecast2_t$	SUE_t	CAR_t	$NEWS1_t$	$NEWS2_t$	$Size_{t-1}$	BM_{t-1}	R_MOM_{t-1}	Adj. R^2	CONTR
Model 1	4.353	-0.049							-0.239	0.223	-0.004	4.84%	11.16%
	2.33	*-2.8*							*-2.32*	*1.37*	*-0.79*		
Model 2	5.226	-0.035	0.019		*0.74*		*1.416*		-0.301	0.303	-0.01	7.81%	5.25%
	2.8	*-1.86*	*9.62*		*9.51*		*16.2*		*-2.63*	*1.92*	*-1.7*		
Model 3	4.122	-0.025		*0.186*		*4.572*		*0.094*	-0.232	0.201	-0.006	24.01%	1.58%
	2.37	*-1.48*		*4.53*		*44.04*		*12.11*	*-2.5*	*1.35*	*-1.28*		

Table 10. Short interest as a predictor of future returns and the role of public news releases: A regression approach

The sample contains common stocks listed on the NYSE, Amex, and Nasdaq during the period from January 2000 to December 2010. We also require that each stock be included in the RavenPack news database. The table shows monthly Fama-MacBeth regressions using excess stock returns over month *t*, calculated as the difference between raw return and one-month T-bill rate, as the dependent variable. We infer news content from the sentiment score provided by RavenPack. *NEWS1* is the average content of news measure in month t where news are classified as either negative, neutral, or positive. We assign numerical values of +1 to all positive news items, zero to all neutral items, and –1 to all negative news items and compute a monthly aggregate news measure by averaging the (signed) values of all news items that month. Months with no news items are assigned an aggregate *NEWS1* measure equal to 0. All news measures are orthogonalized with respect to firm size. *News_Dummy* is a variable that equals to 1 if a news event occurs in month t and 0 otherwise. *Neg_NEWS1_Dummy* (*Pos_NEWS1_Dummy*) is a variable that equals to 1 if a news event occurs in month t and the *NEWS1* content is below (above) median and 0 otherwise. Other independent variables are defined in Table 3. The table reports time-series averages of the coefficient estimates. T-statistics are shown in italics below each coefficient and are based on robust standard errors.

	Dependent Variable: Excess $\textbf{\textit{Ret}}_t$		
Intercept$_{t-1}$	0.022	0.022	0.022
	1.59	*1.63*	*1.64*
Short $_{t-1}$	-0.045	-0.027	-0.027
	-2.26	*-1.27*	*-1.27*
Short$_{t-1}$****News_Dummy***$_t$		-0.050	
		-3.98	
News_Dummy$_t$		0.2	
		1.33	
Neg_ NEWS1_Dummy$_t$			-1.8
			-8.56
Pos_ NEWS1_Dummy$_t$			1.3
			9.77
Short $_{t-1}$***** ***Neg_ NEWS1_Dummy***$_t$			-0.092
			-4.74
Short $_{t-1}$***** ***Pos_ NEWS1_Dummy***$_t$			0.006
			0.43
Size $_{t-1}$	-0.001	-0.001	-0.001
	-0.95	*-1.14*	*-1.13*
BM $_{t-1}$	0.32	0.317	0.366
	2.05	*2.04*	*2.44*
R_MOM $_{t-1}$	-0.002	-0.002	-0.003
	-0.49	*-0.50*	*-0.93*
Adj. R^2	3.90%	4.09%	5.18%

Figure 1. Short interest around earnings announcements and analysts' earnings forecast changes

The sample contains common stocks listed on the NYSE, Amex, and Nasdaq during the period from January 1988 to December 2010. Each month, we sort stocks into quintiles based on four different events and measure the difference in short interest (in percentage) between good news and bad news portfolios starting from 12 months before the event until 12 months after the event. The four different events are defined as follows: (1) standardized unexpected earnings, (*SUE*), defined as the change in earnings per share from quarter *q*-4 to quarter q divided by the standard deviation of unexpected earnings over the last eight quarters, (2) *CAR*s, defined as the average (across firms) of the daily market-adjusted cumulative abnormal return (in percent) during the [-2, +2] window around an earnings announcement date, (3) percentage monthly change in the consensus EPS forecast of analysts defined as $\Delta EPS_t/|EPS_{t-1}|$,and (4) monthly change in consensus EPS forecast of analysts normalized by share price defined as $\Delta EPS_t/P_{t-1}$.

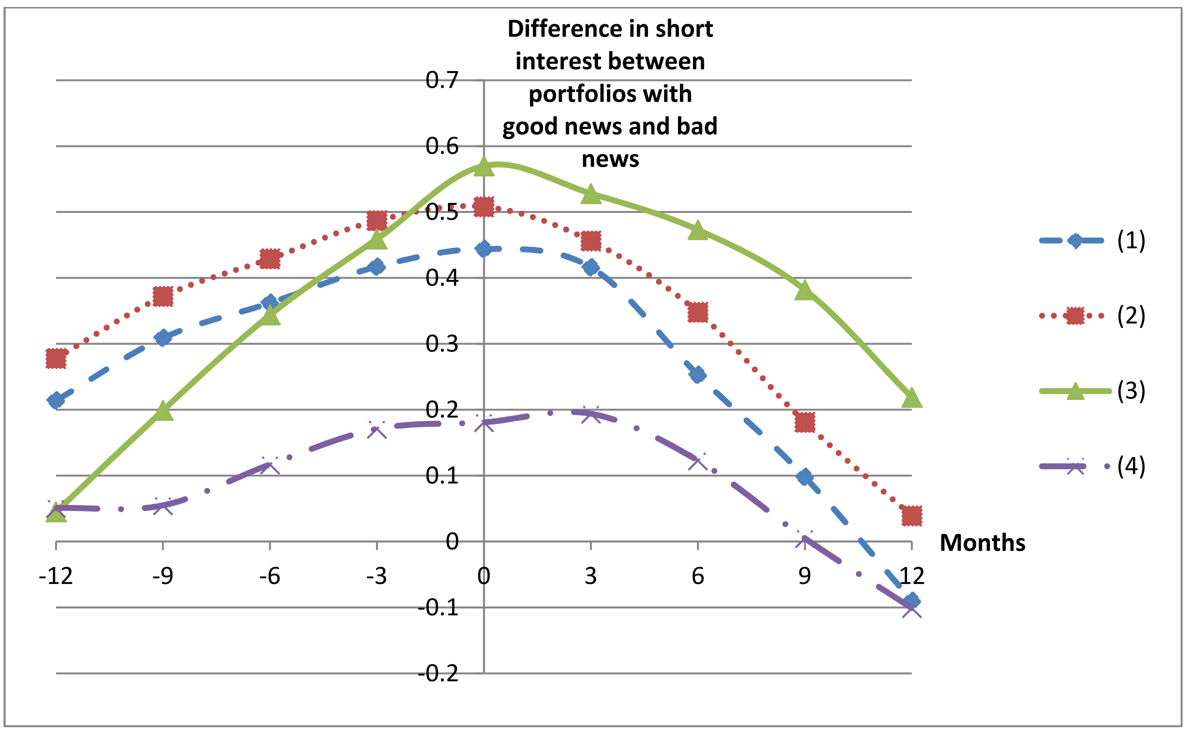

Figure 2. Short interest around public news

The sample contains common stocks listed on the NYSE, Amex, and Nasdaq during the period from January 2000 to December 2010. We also require that each stock be included in the RavenPack news database. Each month, we sort stocks into deciles based on two different new categories and measure the difference in short interest between good news and bad news portfolios starting from 12 months before the event until 12 months after the event. The figure shows the time series averages of the differences. We infer news content from the sentiment score provided by RavenPack. *NEWS1* is the average content of news measure where news are classified as either negative, neutral, or positive. We assign numerical values of +1 to all positive news items, zero to all neutral items, and −1 to all negative news items, and compute a monthly aggregate news measure by averaging the (signed) values of all news items that month. *NEWS2* is the average of sentiment score of news in a month. Months with no news items are assigned an aggregate *NEWS1* (*NEWS2*) measure equal to zero (50). All news measures are orthogonalized with respect to firm size.