Are institutions informed about news? [☆]

Terrence Hendershott [a,*], Dmitry Livdan [a,1], Norman Schürhoff [b,2]

[a] Haas School of Business, University of California, Berkeley, CA 94720, United States
[b] Faculty of Business and Economics at University of Lausanne, Swiss Finance Institute, and CEPR, Extranef 239, CH-1015 Lausanne, Switzerland

ABSTRACT

This paper combines daily buy and sell institutional trading volume with all news announcements from Reuters. Using institutional order flow (buy volume minus sell volume) we find a variety of evidence that institutions are informed. Institutional trading volume predicts the occurrence of news announcements. Institutional order flow predicts (i) the sentiment of the news; (ii) the stock market reaction on news announcement days; (iii) the stock market reaction on crisis news days; and (iv) earnings announcement surprises. These results suggest that significant price discovery related to news stories occurs through institutional trading prior to the news announcement date.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Institutional trading is important because it constitutes the majority of daily trading volume and institutional investors are the largest owners of publicly traded stocks in the U.S.[3] Potentially important drivers of institutional trading are superior information gathering and processing skills. Superior information by institutions could arise from access to more information and greater resources to process information. Unlike retail investors, institutions often directly communicate with publicly traded firms as well as brokerage firms through their investment banking, lending, and asset management divisions. Most mutual funds and hedge funds employ buy-side analysts and enjoy better relationships with sell-side analysts. Their economies of scale allow institutions to monitor many sources of

[3] See, for example, Boehmer and Kelley (2009) and Securities Industry Association Fact Book (2007).

information. Finally, institutions employ professionals and technologies with superior information processing skills. There is some evidence that institutional investors are informed, but studies examining institutional order flow around specific events provide mixed evidence.[4] This paper uses comprehensive news and institutional trading data to show that institutions are informed about news.

To illustrate how institutions trade around news we examine one of the highest profile events in our sample: the Martha Stewart insider trading trial. Stewart's broker tipped her that drug manufacturer ImClone's stock price was about to drop because its drug Erbitux failed to get the expected Food and Drug Administration (FDA) approval. In response, Stewart sold about $230,000 in ImClone shares on December 27, 2001, a day before the announcement of the FDA decision. On June 4, 2003, a federal grand jury in Manhattan indicted Stewart on charges of securities fraud, obstruction of justice, and conspiracy. The same day, Stewart resigned as chief executive officer and chairman of Martha Stewart Living Omnimedia (MSO), but remained on the company's board. Stewart's trial began on January 27, 2004 in New York City and ended on March 5.

Fig. 1 plots MSO's cumulative stock return (Panel A) and institutional order flow (buy volume minus sell volume) in MSO (Panel B) from January 27, 2004 through April 10, 2004. Until February 27, institutions roughly maintained their positions in MSO as their order flow remained close to zero. On February 27, the judge threw out the securities fraud charge against Stewart, which could have led to up to ten years in prison and a $1 million fine. In response MSO's stock price rose roughly 10% and remained there until the verdict on the remaining charges was announced on March 5. In contrast to the rising stock price, institutions sold MSO heavily from February 27 through March 5. Prior to the verdict, institutions sold 8% of MSO's market capitalization.

Trading in MSO was halted after Stewart was found guilty of conspiracy, obstruction of justice, and two counts of making false statements to a federal investigator. When trading in MSO reopened, the stock price plunged roughly 30%. On the same day institutions sold 10% more of MSO's market capitalization. Thus, approximately half of institutions' selling occurred prior to the news. Institutions' selling is consistent with them being better informed about the final verdict and them correctly interpreting the lack of good news in the charges being dismissed on February 27.

Moving beyond the Martha Stewart example to examine whether institutional trading is informed about news in general, this paper combines daily non-public data on buy and sell volume by institutions from 2003 through 2005 for 1,700 NYSE-listed stocks with all news announcements from Reuters. Natural language processing categorizes the sentiment associated with each news story. We use institutional order flow as a quantitative measure of net trading by institutions. Using these comprehensive data of institutional trading and news announcements we find that institutional trading predicts news announcements, the

[4] For example, see the below discussion of Griffin, Shu, and Topaloglu (2012), Jegadeesh and Tang (2010), and Busse, Green, and Jegadeesh (2012).
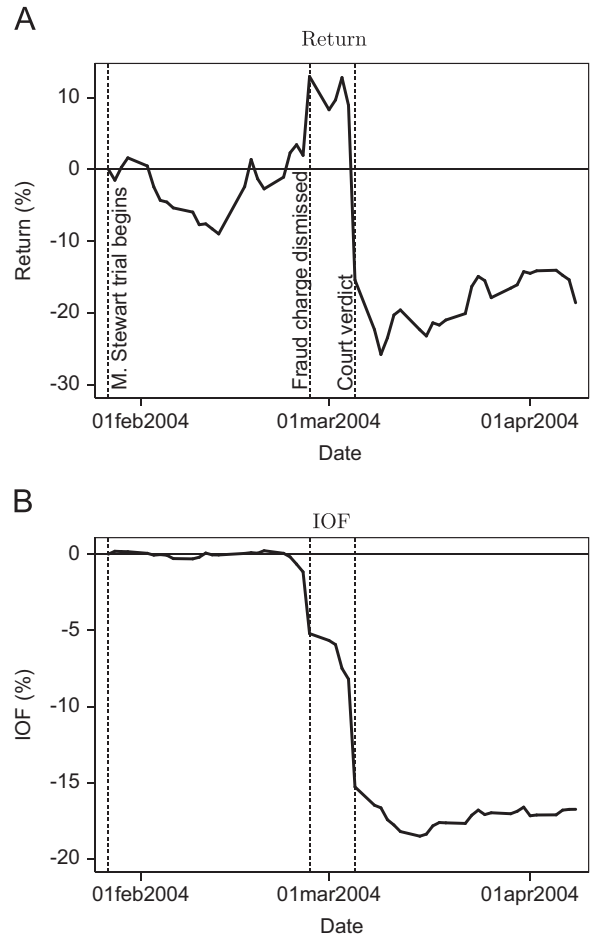
Fig. 1. Stock returns and institutional order flow during the Martha Stewart insider trading trial. The figure documents cumulative stock returns (Panel A) and institutional order flow (IOF; Panel B) in Martha Stewart Living Omnimedia (MSO) from January 27, 2004 through April 10, 2004.

sentiment of the news, returns on the announcement day, and earnings announcement surprises.

To initially examine the question of whether institutions are informed about news, we study institutional trading volume around news announcements (Section 3). Event-study methodology shows that institutional trading volume increases a few days before news announcements. Calendar-time probit regressions show that institutional trading volume predicts whether or not a news announcement will occur after controlling for prior stock volatility and prior news announcements. This is consistent with institutions being informed about whether or not news announcements will occur, although it does not establish that institutions are informed about the content of the news itself.

We next analyze whether institutions are informed about the contents of the news (Section 4). We measure institutions' forecast of future information arrival by their order flow. Natural language processing measures the contents of the news itself. We use stock market reaction on news days as a signal of the information contained in the news announcements. Event-study methodology shows that institutional order flow increases more than five days prior to the announcement of good news as measured by

the natural language sentiment of the news; institutional order flow decreases more than five days prior to bad news announcements. Multivariate regressions show that institutional order flow predicts the sentiment of news announcements and the stock return on announcement days after controlling for prior stock returns, news sentiment, and trading volume. Vector autoregressions that control for longer and more complex joint dynamics of returns, institutional order flow, and news sentiment confirm these results. The economic magnitude of the predictability is meaningful. Applying the Campbell and Thompson (2008) framework implies that observing institutional order flow would enable investors in individual stocks to proportionally increase their expected returns by more than 40%.

Exploiting the Reuters assignment of news stories to various news categories, we investigate the types of news about which institutions are informed (Section 5). First, because asset prices behave very differently on days when important macroeconomic news is scheduled for announcement relative to other trading days (Savor and Wilson, 2014), we investigate institutional trading around macroeconomic news. We find that while institutional trading predicts returns on macroeconomic news days, institutions trade in the direction of macroeconomic news for only one category of macroeconomic news: news on economic indicators.

The 2003–2005 sample is a calm period and institutions' role in firm governance is particularly important during times of stress (Holmstrom and Tirole, 1993). Therefore, we study unexpected value-destroying events. We classify events such as bankruptcy and court decisions as crises. News about these events is associated with negative sentiment and negative stock returns. Institutional order flow predicts returns and sentiment for these categories of bad news.

A potential concern with the prior analysis is that communication between institutions and reporters could affect the sentiment of news. Earnings announcements allow us to study whether institutions are informed about longer term fundamental information, which is very difficult for institutions to affect. We find that institutional order flow predicts the surprise component of earnings announcements.

Finally, we study whether institutions trade in advance of news unrelated to longer term fundamentals, which we refer to as hype. We use two alternative proxies for hype: press releases and news with large subsequent sentiment reversal (i.e., news which is "wrong" ex post). For press releases we link institutional trading with comprehensive news data from PR Newswire and Business Wire. The sentiment of these press releases has a weak correlation of 0.03 with stock returns as compared to a correlation between sentiment and returns of 0.12 for other news stories. We use the Reuters written news data to construct the second hype proxy. Using both proxies we find both qualitatively and quantitatively that institutional order flow shows no abnormal activity around these events. Thus, we find no evidence that institutions trade on news containing little information on fundamentals.

Several studies provide support to the notion that institutions are informed. Badrinath, Kale, and Noe (1995) show that returns of stocks with high institutional ownership lead returns of stocks with low institutional ownership. Sias and Starks (1997) and Boehmer and Kelley (2009) show that higher institutional holdings are associated with more efficient pricing. Boehmer and Wu (2008) and Boulatov, Hendershott, and Livdan (2013) find that institutional trading predicts returns at the firm, industry, and market levels. Irvine, Lipson, and Puckett (2007) find a significant increase in institutional trading and profitable buying beginning five days prior to the public release of analysts' initial reports containing positive recommendations.

Campbell, Ramadorai, and Schwartz (2009) infer institutional trading by linking quarterly changes in institutional holding from 13-F filings with daily trades by size category and a buy–sell classification algorithm. Their measure of institutional trading predicts firms' earnings surprises as does our institutional order flow measure. Campbell, Ramadorai, and Schwartz's (CRS) institutional order flow measure differs in some ways from ours on non-announcement days. The low-frequency component (half-life of 25 days) positively predicts returns whereas the high-frequency component (half-life of one day) negatively predicts returns. These are consistent with institutional trading having a permanent and a temporary price impact. Our measure of institutional trading has only a permanent price impact. This difference is possibly due to CRS needing to infer institutional trading. Overall, we extend CRS' more fundamental point of institutions being informed about earnings news to other types of news.

In contrast, other studies of institutional trading around specific public news events such as takeovers, earnings announcements, and research recommendations find little or no evidence that institutions are informed. Griffin, Shu, and Topaloglu (2012) use Nasdaq broker identifiers on trades and clearing records to categorize trades likely made by institutions from 1997 to 2002. They examine daily trading by eight types of individual and institutional investors ahead of the most common stock market events associated with information asymmetry: takeover and earnings announcements. They find that in the two, five, and ten days prior to takeover announcements, general institutional investors are not net buyers in target firms and their buying is not related to future earnings announcement returns. They do report that hedge funds and investors trading through the largest investment banks that service hedge funds are consistently selling stocks prior to negative earnings announcements. Finally, they find little evidence that brokerage houses' proprietary trading desks or their clients buy prior to takeovers or trade in the right direction prior to earnings announcements.

Jegadeesh and Tang (2010) analyze trading patterns and profitability of institutional trades around takeover announcements using Ancerno's institutional client trade data from 1998 to 2008. They report that institutions on average are marginally net sellers of the targets in the month prior to takeover announcements and that their trading strategy around the announcement does not yield significant abnormal returns. However, they do find that institutions whose main brokers are also the brokerage arms of investment banks advising the targets are significant net buyers of target shares prior to announcements. Using the same data, Busse, Green, and Jegadeesh (2012) examine the performance of buy-side institutional investor trades around sell-side analyst stock recommendations. They find that institutions are

not able to differentiate between good recommendations and bad recommendations.[5]

Our findings along with Campbell, Ramadorai, and Schwartz (2009) suggest that using the broadest possible set of institutional trading data is important to uncover the link between institutional trading and news. As Griffin, Shu, and Topaloglu (2012) discuss, the Ancerno data are less than ten percent of the market. Historically, institutions trade five times more in NYSE than in Nasdaq stocks (Chan and Lakonishok, 1997). This higher institutional activity could explain why Campbell, Ramadorai, and Schwartz (2009) and we find evidence of institutions being informed while other studies do not.

Finally, our paper relates to a growing literature on how different market participants respond to public news. Tetlock (2010) tests a theoretical model with asymmetric information and public news. He finds evidence that news resolves asymmetric information: news has a positive impact on volume-induced return momentum and a temporary increase in the correlation between absolute returns and volume, particularly for earnings news and in small and illiquid stocks.

A separate strand of literature studies whether specific types of institutions such as mutual funds have stock-picking skills prior to public news events. Mutual fund data have the advantage of identifying individual funds and managers, but only do so at monthly intervals. The institutional trading literature combines trading across many institutions at higher frequencies. Baker, Litov, Wachter, and Wurgler (2010) examine the earnings announcement returns of stocks that mutual funds hold and trade. They find that the future earnings announcement returns on stocks that funds buy are, on average, higher than the future returns on stocks that they sell. The stocks that funds buy perform significantly better at future earnings announcements than stocks with similar characteristics, while the stocks that funds sell perform significantly worse than matching stocks. Fang, Peress, and Zheng (2014) examine the propensity of mutual funds to trade high media-coverage stocks. They find that funds with a lower propensity to trade high media-coverage stocks perform significantly better. This finding is robust to different risk adjustment models and holds after controlling for other fund characteristics. Their result is consistent with the hypothesis (see Kacperczyk and Seru, 2007) that funds with informational advantage trade less in stocks with media coverage.

Several papers examine the relationship between individual trading and news announcements. Kaniel, Liu, Saar, and Titman (2012) provide evidence in support of informed trading by showing that intense aggregate individual investor buying (selling) predicts large positive (negative) abnormal returns on and after earnings announcement dates. Kelley and Tetlock (2013) use retail brokers' trading data from 2003 to 2007 to provide support for the conclusions of Kaniel, Liu, Saar, and Titman (2012) that retail investors have some information for a broader set of news announcements. Kelley and Tetlock (2013) can separately identify market and limit orders and find that market order imbalances predict both returns and news, whereas limit order imbalances predict returns but not news.

Short selling is another type of trading thought to be informed (Senchack and Starks, 1993; Asquith, Pathak, and Ritter, 2005; Boehmer, Jones, and Zhang, 2008; and others). Engelberg, Reed, and Ringgenberg (2012) combine data on short selling with news releases to show that short sellers' trading advantage comes largely from their ability to analyze publicly available information and not from being able to anticipate information before it becomes public. In contrast, we find that institutions overall are able to anticipate information before it becomes public news.

The remainder of the paper is organized as follows. Section 2 discusses the data sources and provides summary statistics. Section 3 examines institutional trading volume around news announcements. Section 4 analyzes whether institutions are informed about the contents of the news. Section 5 investigates which specific type(s) of news institutions are informed about. Section 6 concludes.

2. Data

The data on trading by institutions are constructed from the NYSE's Consolidated Equity Audit Trail Data (CAUD) files, which contain detailed information on all orders that execute on the exchange. One of the fields associated with the buyer and seller of each order, Account Type, specifies whether the order comes from an institutional investor. We exclude program trading and index arbitrage trading because these order types are for trading multiple securities simultaneously and thus are less likely related to news about individual stocks.[6] We supplement the CAUD data with daily data on returns (close-to-close returns based on closing bid and ask quotes in Trade and Quote (TAQ)), trading volume from the Center for Research in Security Prices (CRSP), and market capitalization (number of shares outstanding times price from CRSP).

Our news data come from the Thomson Reuters News Analytics (TRNA), which is a database of news releases on the Reuters Data Feed (RDF). TRNA uses a neural network to construct measures of news sentiment each news story reports. The analysis primarily is conducted at the sentence level. During the initial pre-processing stage firms mentioned in each sentence are identified and then what is said about these firms is analyzed. Each word in a sentence is parsed into a sentence tree according to the corresponding part of speech. The parts of speech are then fed into a neural network classifier. The neural network was trained on several thousand randomly selected news stories which were tagged by three former traders. The neural network tries to incorporate the order of words, adjectives, and common phrases in finance. Sinha (2012) and Infonic (2008) provide further discussion of TRNA's text processing.

TRNA has several important differences from the popular dictionary-based method for analyzing text.[7] TRNA's text-

---

[6] The TORQ data set provides a sample of the CAUD data.

[7] Tetlock, Saar-Tsechansky, and Macskassy (2008) use dictionary-based methods to count words based on a general Harvard psychosocial dictionary and Loughran and McDonald (2011) use a specialized financial dictionary. Heston and Sinha (2013) compare how well the TRNA sentiment and sentiment calculated from these dictionary approaches forecasts future returns.

processing engine analyzes at the sentence level rather than the word level. Analyzing a document at the sentence level is beneficial along a number of dimensions. The sentence-level analysis ensures that the word is analyzed in its context. Modifiers, e.g., negative construction, adjectives, or adverbs, alter the meaning of words. In addition, firms choose names which dictionaries classify as positive. Every time the story refers to the firm's name, a pure dictionary approach counts a positive word. Finally, the sentence approach identifies the sentence's subject.

When a story mentions multiple firms, each sentence is correctly attributed to the corresponding firm. This is used to determine different sentiment for different firms in the same story. The relevance of the story for each firm is based on comparing the number of mentions of a firm to the total number of mentions of all firms in the story.

Each news story on the RDF typically consists of several news items. Each day we average the sentiment for each story and then construct a daily weighted average sentiment across stories using the relevance measure as weights. News items are either a new alert or new story take, in which an alert is a single line of text and a story take has a headline and body. A story take is one in a series of updates to a particular story. Alternative schemes to aggregate news items to stories and, respectively, the stories on a news day do not materially affect the results. To align the story dates with prices, returns, and institutional trading, the date associated with each story is set using a cutoff of the NYSE closing time of 4 pm Eastern time. Stories appearing after 4 pm are given the following date.

Table 1 provides an overview of the different news topics covered by the RDF in descending order based on the total number of news days reported in the fourth column. Statistics for the relevance-weighted news sentiment are in columns five through nine. Reuters assigns stories to news categories describing the topic of the news content. The most frequent topics are directly related to companies and concern corporate results forecasts (RESF) with a total of 33,409 news days, corporate results (RES, 32,054), corporate crises (CRISIS, 31,675), debt markets (DBT, 28,214), stock markets (STX, 27,543), major breaking news (NEWS, 26,639), corporate bonds (USC, 25,091), mergers and acquisitions (MRG, 23,791), macro news (MACRO, 17,614), business activities (BACT, 16,786), corporate analysis (CORA, 16,168), hot stocks (HOT, 14,541), regulation (REGS, 14,294), government policies (WASH, 13,097), legislation (LAW, 11,217), fund industry news (FUND, 11,176), broker research and recommendations (RCH, 10,209), ratings (AAA, 9,549), new issues (ISU, 9,406), job losses and unemployment (JOB, 8,776), and management issues and policy (MNGISS, 7,346). The remaining news releases comprise specific macroeconomic announcements, government policies and politics, society, environment, and other financial market news.

We construct our sample by merging the CRSP, TRNA, and NYSE data from 2003 through 2005 and dropping a small number of observations (0.51%) for which some data are missing from one of the sources. There are a total of 755 trading days in 1,667 stocks, yielding more than one million daily observations with complete data on stock return, trading volume, news, institutional trading, and all control variables. Table 2, Panel A provides summary statistics for the number of news releases and the distribution of news stories across time and stocks. There are a total of 126,148 days with news releases out of 1,096,514 daily observations during our sample period. This implies that 11.5% of stocks have news releases on any given day, with 5.9% (17.6%) of stocks in the news at the lower (upper) 5% tail of days. Consistent with prior papers using news data, there is substantial cross-sectional variation in news coverage. The average firm has a 10.6% chance of being covered in a news report. While the median firm has a propensity of news coverage of once per month (4.6%), news coverage ranges from zero for the bottom 5% of firms to 47.5% for the top 5% of firms.

Individual news stories can be relevant for one or multiple firms. In order to distinguish firm-specific from sectoral news, news items read from the RDF are scored with respect to companies that are mentioned in the article to yield company-specific measures of relevance. The third row in Table 2, Panel A shows that the average (median) number of firms mentioned in a news release is 7.2 (4).

The news sentiment measure used in this study is based on the analysis of the NewsScope news text released on the RDF. The Reuters algorithm determines how positive, neutral, or negative is the tone of the words used in the article for each firm. Individual sentiment scores yield the positive, neutral, or negative sentiment score for the news item, ranging between -1 and 1.

We compute the net sentiment of a news story as the relevance-weighted difference between the positive and negative score for each news item. We then aggregate all news stories on a given day by relevance weighting the story-specific sentiment to obtain the daily sentiment in each stock. The net sentiment is set to zero on days without news stories. The last row in Table 2, Panel A provides summary statistics for daily sentiment. Sentiment ranges between -0.726 and 0.738 at 5% and 95%, respectively, with mean and median close to zero and a standard deviation of 0.421.

Institutional purchases, $IBuys_{i,t}$, and sales, $ISales_{i,t}$, aggregate all institutional buy and, correspondingly, sell transactions for a firm i on day t. These quantities are then normalized by the firm's market capitalization, MC, lagged by one year, yielding

$$IBuys_{i,t} = \frac{\sum_{n=1}^{\text{Number of Buys}_{i,t}} Buys_{i,t}^n}{MC_{i,t-250}},$$

$$ISales_{i,t} = \frac{\sum_{n=1}^{\text{Number of Sales}_{i,t}} Sales_{i,t}^n}{MC_{i,t-250}}. \quad (1)$$

Institutional order flow, IOF, is the difference between institutional purchases, IBuys, and institutional sales, ISales. Institutional volume, IVol, is the sum of institutional purchases and sales.

Table 2, Panel B provides summary statistics for the institutional trading volume and order flow imbalances across stocks in our sample. Institutional order flow imbalances are positive on average, consistent with the steady decline in direct individual stock ownership over time. Institutional order flow imbalances are distributed symmetrically around this mean, with significant negative and positive days. This shows that despite the positive

Table 1
Description of news categories.

The table provides a brief description of the major news categories related to individual stocks in the TRNA database. The sample contains all 1,667 NYSE stocks on news days from 2003 through 2005.

Topic	Topic code	Description	News days	Sentiment				
				Mean	SD	5%	50%	95%
Corporate Results Forecasts	RESF	All forecasting of corporate financial results	33,409	−0.035	0.463	−0.754	−0.011	0.726
Corporate Results	RES	All corporate financial results; dividends, accounts, and annual reports	32,054	−0.012	0.433	−0.744	0.036	0.679
Corporate Crisis	CRISIS	All corporate crisis (composite of BKRT, MNGISS, JOB, CRIM, JUDIC, REGS, CDV, WEA)	31,675	−0.221	0.436	−0.762	−0.335	0.615
Debt Markets	DBT	All debt market news, including primary issuance, trading, market forecasts, and analysis	28,214	−0.025	0.463	−0.754	0.018	0.752
Stock Markets	STX	All news about equity markets operations, regulations and structure, etc.	27,543	−0.002	0.331	−0.705	0.040	0.620
Major Breaking News	NEWS	Top stories of major international impact (likely to lead TV/ radio/newspaper bulletins)	26,639	−0.134	0.475	−0.759	−0.222	0.711
US Corporate Bonds	USC	All news about US corporate bonds, including issues, forecasts, and analysis	25,091	−0.033	0.477	−0.756	0.004	0.756
Mergers and Acquisitions	MRG	All corporate stories about change of ownership; stakes, mergers, acquisitions, buy-outs	23,791	0.197	0.474	−0.674	0.271	0.808
Macro News	MACRO	All macro news (composite of ECI, FED, GVD, MCE, WASH)	17,614	−0.189	0.441	−0.761	−0.271	0.653
Business Activities	BACT	News relating to business activities	16,786	0.038	0.101	0.032	0.040	0.044
Corporate Analysis	CORA	Analysis about a company or group of companies	16,168	0.039	0.059	0.032	0.040	0.043
Hot Stocks	HOT	News about stocks "on the move"	14,541	−0.102	0.528	−0.762	−0.201	0.779
Regulatory Issues	REGS	News about regulation	14,294	−0.261	0.431	−0.762	−0.400	0.596
Washington/US Govt. News	WASH	All US Federal government politics, policies, and economics	13,097	−0.232	0.432	−0.762	−0.352	0.595
Legislation	LAW	Legislation affecting budgets, securities laws, capital budgets, lawsuits, court rulings	11,217	−0.399	0.373	−0.763	−0.504	0.399
Fund Industry News	FUND	All news about investment trusts and funds industry; funds' views and forecasts	11,176	−0.029	0.491	−0.760	0.007	0.776
Broker Research and Recom.	RCH	All news about broker research and recommendations	10,209	0.008	0.579	−0.763	0.063	0.812
Ratings	AAA	All news about credit ratings	9,549	0.007	0.512	−0.758	0.059	0.792
New Issues	ISU	All new government and corporate issues of debt and corporate issues of equity	9,406	0.113	0.381	−0.572	0.114	0.742
Labor; (Un) employment	JOB	All news on labor issues; (un)employment, labor disputes, strikes, unions, etc.	8,776	−0.153	0.470	−0.761	−0.238	0.724
Management Issues/Policy	MNGISS	Management issues including executive pay, bonuses, governance, accounting irregularity	7,346	−0.161	0.440	−0.762	−0.173	0.595
Domestic Politics	POL	All stories about national politics	6,481	−0.231	0.443	−0.762	−0.376	0.644
Mortgage Backed Debt	MTG	Asset-backed securities, mortgage-backed debt, and changes in mortgage rates	6,468	0.007	0.456	−0.753	0.041	0.772
Internet/World Wide Web	WWW	All news stories relating to the Internet or the World Wide Web	6,284	−0.035	0.489	−0.757	−0.004	0.762
Reuters Exclusive News	WIN	Major news, exclusive to Reuters	5,463	0.029	0.477	−0.742	0.059	0.772
Derivatives	DRV	All news and market reports on derivatives, including futures, options, and swaps	5,452	−0.090	0.450	−0.756	−0.086	0.730
Press Digests	PRESS	Summaries of newspaper articles	4,882	−0.214	0.484	−0.763	−0.407	0.763
Multi-Industry	MUL	All news on diversified companies, including holding companies	4,581	−0.018	0.498	−0.761	0.033	0.777
Loans	LOA	All types of loans to corporate entities and sovereign countries	4,418	−0.013	0.424	−0.751	0.061	0.702
Crime, Law Enforcement	CRIM	Civil and criminal law, corporate crime, fraud, murder, criminals, mafia, etc.	3,928	−0.460	0.347	−0.763	−0.531	0.342
Terms of Bond Issues	TNC	News about bond terms and conditions	3,844	0.084	0.285	−0.460	0.087	0.515
Interest Rates	INT	All news on interest rates and interest rate changes and forecasts; analysis of rate moves	3,814	−0.045	0.446	−0.741	−0.032	0.767
Government/ Sovereign Debt	GVD	All government debt market news, government borrowing, and debt	3,689	−0.103	0.420	−0.754	−0.093	0.692
Dividends	DIV	Dividend forecasts, declarations, and payments	3,631	0.293	0.472	−0.718	0.449	0.811
Judicial Processes/ Court Cases	JUDIC	All stories about judicial processes/court cases/court decisions	3,408	−0.459	0.335	−0.763	−0.512	0.282
Initial Public Offerings	IPO	First public listing of a company's stock	3,329	0.150	0.432	−0.638	0.189	0.793
Eurobonds	EUB	New eurobond issues, including issues by foreign borrowers in domestic markets	2,771	0.013	0.401	−0.730	0.070	0.707
	CDV	All news relating to credit default swaps	2,467	−0.195	0.496	−0.762	−0.324	0.745

Table 1 (*continued*)

Topic	Topic code	Description	News days	Sentiment				
				Mean	SD	5%	50%	95%
Credit Default Swaps								
Wholesale	WHO	Wholesaling and distribution, including wholesale price indices	2,390	−0.017	0.472	−0.759	0.045	0.775
International Trade	TRD	All stories associated with international trade; protectionism, tariffs, sanctions, dumping	2,343	−0.165	0.482	−0.761	−0.284	0.726
Weather	WEA	All weather issues and reports; forecasts, statistics, warnings, etc.	2,236	−0.262	0.407	−0.760	−0.383	0.531
Bankruptcies	BKRT	Corporate insolvencies and bankruptcies, creditor protection actions, court rulings, etc.	2,070	−0.353	0.416	−0.763	−0.497	0.538
Macro-Economics	MCE	All news and analysis on macroeconomics	1,918	−0.110	0.503	−0.761	−0.192	0.791
Economic Indicators	ECI	News, forecasts, or analysis of economic indicators	1,814	−0.127	0.493	−0.762	−0.223	0.774
US Agencies	AGN	All news about debt-issuing agencies such as Fannie Mae, Freddie Mac, etc.	1,810	−0.114	0.300	−0.634	−0.093	0.372
Forex Markets	FRX	All market stories about foreign exchange, forex intervention by central banks, etc.	1,738	−0.054	0.501	−0.760	−0.050	0.776
Investment Grade Debt	IGD	News on debt classed as investment grade	1,642	−0.089	0.500	−0.761	−0.084	0.779
Diplomacy, Int. Relations	DIP	Political relations between multiple countries; foreign policy by governments	1,618	−0.210	0.462	−0.762	−0.362	0.712
Money Markets	MMT	All money market news; reports about money markets	1,569	−0.089	0.319	−0.630	−0.054	0.414
Disasters and Accidents	DIS	Natural and man-made disasters; major industrial accidents, pollution disasters, etc.	1,512	−0.300	0.393	−0.762	−0.443	0.425
Exchange Activities	EXCA	Stories about exchanges where securities or futures trading takes place	1,312	−0.061	0.467	−0.760	−0.009	0.733
Lifestyle	LIF	All news about lifestyles; lifestyles of people in the public eye	1,302	−0.079	0.439	−0.745	−0.048	0.680
Civil Unrest	VIO	Stories about riots, demonstrations, and other internal disturbances	1,148	−0.354	0.410	−0.763	−0.498	0.528
High-Yield Debt	HYD	All news about bonds rated below BBB- or ratings actions on issuers below BBB-	1,059	−0.173	0.528	−0.763	−0.322	0.774
Federal Reserve Board	FED	FED activities and news	990	−0.197	0.453	−0.762	−0.269	0.677
(Inter)national Security	SECUR	All stories about national and international security	902	−0.216	0.470	−0.762	−0.404	0.727
Short-Term Interest Rates	STIR	Short-term interest rates, typically for maturities up to two years	812	−0.187	0.412	−0.706	−0.294	0.765
Equity-Linked Bonds	EQB	Equity-linked bonds, warrant bonds, or convertible bonds	745	0.104	0.387	−0.683	0.088	0.744
Asset-Backed Debt	ABS	All news about asset-backed debt, including credit card receivables and auto loans	698	0.159	0.460	−0.743	0.281	0.788
Reuters Summits	RSUM	News from Reuters industry summits	656	0.076	0.467	−0.703	0.086	0.798
Human Interest	ODD	Unusual, offbeat, and curious stories	548	−0.151	0.442	−0.757	−0.237	0.651
Muni and Agency Debt	REVS	News about debt sold by state and local authorities and agencies	489	−0.082	0.508	−0.763	−0.022	0.787
Hedge Funds	HEDGE	News about private investment funds	313	0.059	0.533	−0.763	0.132	0.806
Elections	VOTE	National, regional, local elections; manifestos, polling (relates only to government)	254	−0.266	0.445	−0.763	−0.455	0.670
Tax	TAX	All news about taxation rules and regulation, especially relating to individuals	212	−0.145	0.492	−0.760	−0.291	0.767
Religion	REL	All matters relating to religion and religious institutions	180	−0.246	0.433	−0.757	−0.439	0.566
Bombings	BOMB	All stories about bombings	67	−0.490	0.327	−0.763	−0.530	0.275
Fiscal and Monetary Policy	PLCY	Statements by or stories about fiscal and monetary policy makers	67	−0.290	0.358	−0.763	−0.319	0.301
Editorial Specials	FES	Editorial special, analysis, and future stories	40	−0.039	0.504	−0.714	−0.122	0.793
Errors	ERR	Errors	30	0.132	0.366	−0.479	0.133	0.675
Investing	INV	All news about the process of investing on the part of individuals	24	0.191	0.494	−0.514	0.261	0.807
Retirement	RTM	All news about retirement, including financial regulation and industry trends	13	−0.295	0.422	−0.761	−0.493	0.538
Technical Analysis	INSI	Stories about technical analysis of markets	8	0.066	0.477	−0.689	0.153	0.533
Corporate Finance	CFIN	All news relating to corporate finance	5	−0.142	0.472	−0.660	−0.016	0.368
Emerging Markets	EM	All news relating to emerging markets	4	0.085	0.602	−0.764	0.261	0.584
Purchasing Managers Indices	PMI	News coverage of any Purchasing Managers' Index of manufacturing and/or services	1	−0.478	0.000	−0.478	−0.478	−0.478

Table 2
Descriptive statistics.

The table reports descriptive statistics for the news data and the institutional trading data in our sample. News are aggregated by stock day. Sentiment is computed as the relevance-weighted average of the difference between positive and negative sentiment scores. Institutional order flow IOF (institutional volume IVol) is defined as the difference between (sum of) institutional purchases IBuys and institutional sales ISales. All trade-related quantities are normalized by the firm's market capitalization lagged by one year and expressed in percent. The sample contains all 1,667 NYSE stocks on all 755 trading days from 2003 through 2005. Levels of significance are denoted by * (10%), ** (5%), and *** (1%).

	Mean	S.D.	5%	50%	95%
Panel A: News releases and sentiment (126,148 observations)					
News stocks per day (% of 1,667 firms)	11.506	3.921	5.886	11.187	17.595
News days per stock (% of 755 days)	10.610	16.667	0.000	4.636	47.458
Stocks per news release	7.197	9.729	1.000	4.000	24.000
Sentiment per news release	0.007	0.421	−0.726	0.040	0.738
Panel B: Institutional trading (1,096,514 observations)					
IOF (% of size)	0.004	0.168	−0.176	0.002	0.189
IVol (% of size)	0.832	1.644	0.046	0.432	2.728
|IOF|/IVol	0.153	0.166	0.008	0.102	0.475
IBuys (% of size)	0.418	0.829	0.020	0.216	1.375
ISales (% of size)	0.414	0.824	0.019	0.213	1.367
Panel C: Return and volume (1,096,514 observations)					
Return (%)	0.091	1.945	−2.913	0.034	3.241
Volume (% of size)	1.767	3.205	0.144	1.002	5.479

Panel D: Correlations	$Sentiment_t$	IOF_t	$IVol_t$	$Return_{t-1}$	$Sentiment_{t-1}$	IOF_{t-1}
$Return_t$ on news days	0.122***	0.086***	0.015***	0.007**	0.011***	0.022***
$Return_t$ on non-news days	–	0.052***	0.043***	−0.006***	0.028***	0.014***

trend there are about the same portion of days when institutions are net sellers as when they are net buyers. Institutional order imbalances are small compared to overall institutional trading activity. IVol shows that on a typical day, 83.2 basis points of stock market capitalization are traded by institutions while only 15.3% of this is net purchases or sales from other investor groups, such as retail investors, market makers, and institutions trading baskets of stocks.

Panel C of Table 2 reports summary statistics for stock returns and market-wide trading volume. Comparing IVol to turnover shows that IVol is roughly 47% of CRSP trading volume in our sample. On news days both sentiment and returns and IOF and returns are positively contemporaneously correlated: the correlation coefficients are 0.122 and 0.086, respectively. Finally, the return correlation is higher on news days than on non-news days for IOF, but lower for IVol.

3. Institutional trading around news releases

First we test whether institutions adjust their overall trading activity ahead of future public news announcements and whether institutional trading predicts news announcements. Fig. 2 graphically demonstrates this relation by plotting results from an event study. Panel A depicts institutional trading volume, IVol, in the [−10, 10] window around the news announcement. Dashed lines represent standard errors adjusted for heteroskedasticity, contemporaneous correlation across stocks, and autocorrelation within each stock by clustering on day and stock throughout the analysis as

suggested by Petersen (2009). We use market value weighting throughout. The baseline value for IVol in Fig. 2, Panel A is lower than the equal-weighted average reported in Table 2, Panel B since IVol correlates negatively with market capitalization.

Institutional trading volume rises sharply before the news announcement day, from a steady value-weighted average of about 0.43% to 0.48% of a firm's market capitalization, and it declines sharply after the news has become public. These patterns are consistent with the hypothesis that institutions are privately informed about future public news. An alternative story is that the rise in institutional trading leads to higher return volatility which in turn gets noticed by the news agencies which respond with news articles. Or simply put, news agencies track actively traded and volatile stocks and write news stories about them. We examine this hypothesis in Panel B of Fig. 2 by plotting the absolute stock returns, |Return|, which proxy for return volatility, over ten days before and ten days after news announcements. As in the case of institutional trading volume, return volatility rises sharply before the announcement and then sharply declines thereafter. To disentangle these effects we study the joint relations among news announcements, institutional trading, and |Return|.

Table 3, Panel A presents estimates from panel logit regressions with the dependent variable being zero or one depending on whether a news announcement involving firm i takes place on date t. Firm fixed effects are included in the specification to control for the cross-sectional heterogeneity in news announcement frequency. Column A reports the univariate regression with lagged IVol as the
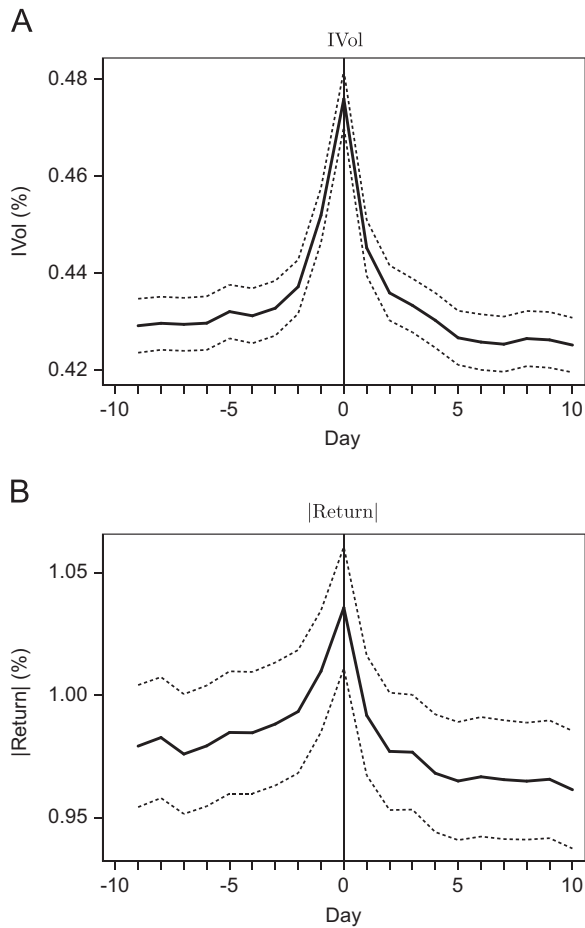
A

B

Fig. 2. Institutional trading volume and stock return volatility around news announcements. The figure documents institutional trading volume and stock return volatility around news announcements. Panel A reports institutional volume, measured as a fraction of market capitalization, between ten days before and ten days after a news announcement. Panel B reports absolute stock returns over the same time period. The mean values are calculated as the value-weighted average of the individual news day values. The dotted lines indicate 95% confidence bounds. The sample contains all 1,667 NYSE stocks on news days from 2003 through 2005. The number of observations is 126,148. Observations are value-weighted. Standard errors are robust to heteroskedasticity and clustering.

explanatory variable. In agreement with the event study from Panel A of Fig. 2, the regression coefficient on IVol is positive and statistically significant. Next, we use |Return| as the explanatory variable. The positive, statistically significant coefficient on |Return| in column B of Table 3 is consistent with the event study in which the return volatility increases prior to news announcements. Columns C and D of Table 3 report results when the explanatory variables are an indicator variable for news announcements on the previous day, Newsday, and the absolute value of the prior news sentiment, |Sentiment|. Column C indicates that news announcements are persistent, i.e., news clusters together in time, as stocks previously in the news are more likely to be in the news again. In addition, the propensity of new news stories increases when the prior news story had more significant sentiment, as the coefficient on the lagged

absolute sentiment is positive in column D. Columns E and F report the results when all four variables are used together as explanatory variables. All of them remain statistically significant in the multivariate regression with the inclusion of day fixed effects in the last column. The point estimates for IVol and |Return| are larger with the inclusion of day fixed effects, which is consistent with institutional trading being more related to idiosyncratic news as opposed to market-wide news.

Panel B of Table 3 uses the number of news stories as a dependent variable to test whether past institutional trading puts stocks into the spotlight. We find that past institutional trading (column A), |Return| (column B), and having recently been in the news (columns C and D) are good predictors of the number of future news articles written. Therefore, firms that are in the spotlight tend to stay in the spotlight for some time and their past market performance puts them under the spotlight in the first place. Overall, the results on institutional trading volume and the occurrence of news announcements are consistent with the hypothesis that institutions have private information about future news, but this is not the only possible interpretation of the evidence.

4. Are institutions informed about public news?

While the previous section provides evidence that trading by institutions is related to future news releases, it does not establish that institutions actually are informed about contents of the news. To address this question we study if institutional buying and selling predicts the sentiment of the news and the stock price reaction to the news. For institutions to be informed about the contents of the news, more buying should predict news announcements with positive sentiment and positive price reactions; similarly, more selling should predict negative news announcements and negative price reactions. As in our analysis of aggregate institutional trading volume and news, we examine event-study and regression evidence.

4.1. Event-time evidence

To investigate the informativeness of institutional trading we examine whether trading predicts the announcement day abnormal return. To do so we calculate buy-and-hold abnormal stock returns (BHAR) for each news release per firm. We also differentiate between different types of news by categorizing them as Good or Bad news. We define Good and Bad news as a function of the sentiment associated with the news release by dividing announcement sentiment into quintiles across the pooled set of announcements. Good are news releases associated with sentiment in the top quintile across all news announcements, Sentiment ≥ 0.374. Correspondingly, Bad news is when the news sentiment is in the bottom quintile, Sentiment ≤ -0.418. The results are not sensitive to the exact cutoff values. All BHARs are benchmarked against a control group of firms.

Table 3

Predicting public news announcements.

The table documents the predictability of public news announcements. The dependent variable indicates a news announcement on date t in firm i (Panel A) or, alternatively, the number of news stories on date t in firm i (Panel B). Estimates are from panel logit (negative binomial) regressions with firm fixed effects in Panel A (B). IVol is institutional volume, |Return| is the absolute daily stock return, and Newsday is the lagged news announcement indicator variable. The sample contains all 1,667 NYSE stocks on all 755 trading days from 2003 through 2005. The number of observations is 1,096,514. Observations are value-weighted. Standard errors are robust to heteroskedasticity and clustering and reported in parentheses. Levels of significance are denoted by * (10%), ** (5%), and *** (1%).

	(A)	(B)	(C)	(D)	(E)	(F)
Panel A: Newsday$_t$						
IVol$_{t-1}$	0.025***				0.017***	0.164***
	(0.000)				(0.000)	(0.000)
|Return$_{t-1}$|		9.655***			1.493***	7.213***
		(0.000)			(0.000)	(0.000)
Newsday$_{t-1}$			0.906***		1.834***	0.776***
			(0.000)		(0.000)	(0.000)
|Sentiment$_{t-1}$|				0.992***	−0.752***	−0.150***
				(0.000)	(0.000)	(0.000)
Day fixed effects	No	No	No	No	No	Yes
Log-likelihood	−2,405	−2,403	−2,343	−2,388	−2,378	−2,104
Panel B: No. of news stories$_t$						
IVol$_{t-1}$	0.063***				0.042***	0.168***
	(0.000)				(0.000)	(0.000)
|Return$_{t-1}$|		6.662***			5.011***	7.684***
		(0.000)			(0.000)	(0.000)
Newsday$_{t-1}$			0.712***		0.803***	0.349***
			(0.000)		(0.000)	(0.000)
|Sentiment$_{t-1}$|				0.409***	−0.331***	−0.012***
				(0.000)	(0.000)	(0.000)
Day fixed effects	No	No	No	No	No	Yes
Log-likelihood	−1,107	−1,107	−1,092	−1,106	−1,090	−1,063

The buy-and-hold return in the $[t_0, t_1]$ window for an event-firm $i \in (Good, Bad)$ is defined as

$$BHR_i(t_0, t_1) = \prod_{t_0}^{t_1} R_{i,t}, \qquad (2)$$

where $R_{i,t}$ is the gross return of firm i on date t. The mean abnormal buy-and-hold returns for good and bad news firms are

$$BHAR(t_0, t_1) = \sum_i w_i BHR_i(t_0, t_1) - \sum_{Control} w_{Control} BHR_{Control}(t_0, t_1), \qquad (3)$$

where w_i is firm i's market capitalization weight lagged by one year divided by the number of events so that $\sum_i w_i = 1$. The mean $BHAR$ is calculated as the value-weighted average of the individual event-firm BHRs benchmarked against the mean BHRs for all control firms. For simplicity, we take the value-weighted index of all firms in our sample as the control. The results are not materially affected when we use Fama-French-Carhart four-factor residuals.

If institutions are privately informed about news prior to the publication date, then their trading can cause significant price run-up/run-down prior to the good/bad news. Alternatively, the large price movements in positive/negative direction could lead to news and its sentiment. Similar to Fig. 2 we calculate buy-and-hold abnormal returns starting at $t_0 = -10$ relative to the news release at date $t=0$ and through the ten trading days following the news. Panel A of Fig. 3 reports our results for average buy-and-hold returns around good and bad news releases. The dotted lines correspond to 95% confidence bounds. Prices

begin to drift in the direction of the news sentiment a few days before announcement day, consistent with models of private information prior to the public news announcement. The largest price run-up/run-down happens in the days immediately prior to the news announcement. This could be because informed traders, possibly institutions, trade more aggressively over time, as in Kyle (1985) and Back, Cao, and Willard (2000), or because traders become more informed about the news as the announcement day approaches.

Cumulative institutional order flows before an announcement provide a measure of institutional trading potentially driven by private information. Analogous to the $BHARs$ in Panel A of Fig. 3, we calculate buy-and-hold institutional order flow for each firm experiencing a news release. $IOF_{i,t}$ is the institutional order flow of firm i on date t. The buy-and-hold institutional order flow in the $[t_0, t_1]$ window for an event-firm $i \in (Good, Bad)$ is defined as

$$BHIOF_i(t_0, t_1) = \sum_{t_0}^{t_1} IOF_{i,t}. \qquad (4)$$

Then similar to returns we compute the mean abnormal buy-and-hold institutional order flow for good and, respectively, bad news firms as in the case of buy-and-hold returns:

$$BHAIOF(t_0, t_1) = \sum_i w_i BHIOF_i(t_0, t_1)$$
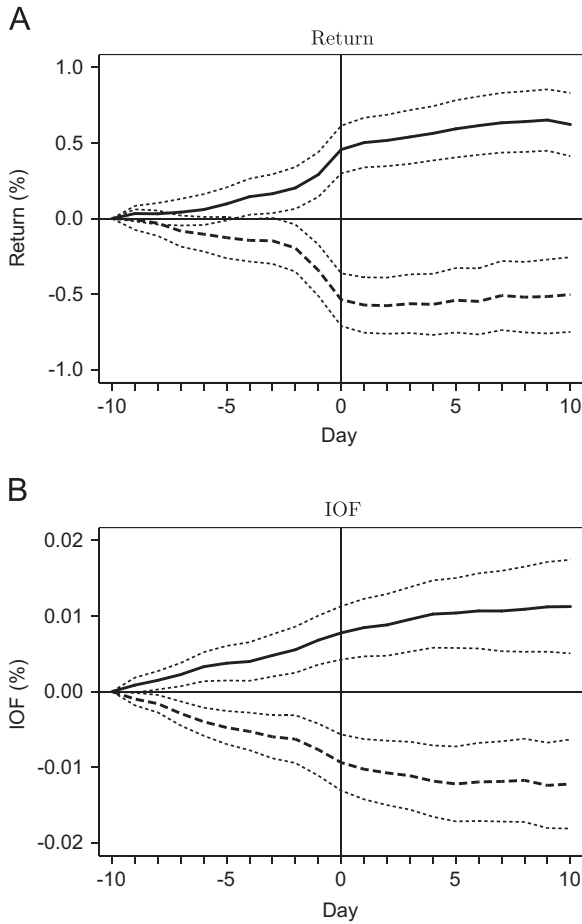$$- \sum_{Control} w_{Control} BHIOF_{Control}(t_0, t_1). \qquad (5)$$

A

B

Fig. 3. Institutional order flow and stock returns around news announcements. The figure documents institutional trading and stock returns around news announcements. Panel A reports buy-and-hold cumulative stock returns between ten days before and ten days after a news announcement. Panel B reports buy-and-hold cumulative institutional order flow over the same time period. The solid line represents good news days, defined as the top sentiment quintile on news announcement days. The dashed line represents bad news days, defined as the bottom sentiment quintile on news announcement days. The mean values are calculated as the value-weighted average of the individual news day values. The dotted lines indicate 95% confidence bounds. The sample contains all 1,667 NYSE stocks on news days from 2003 through 2005. The number of observations is 126,148. Observations are value-weighted. Standard errors are robust to heteroskedasticity and clustering.

As before we use the value-weighted index of all firms in our sample as the control.

Fig. 3, Panel B summarizes the IOF results. The dotted lines correspond to 95% confidence bounds. Institutions start net buying ten days before good news announcements and net selling ten days before bad news announcements. The order imbalances are largest during the two days prior to the announcement, corresponding to the larger returns on these days.

The event study in Fig. 3 considers only IOF and news. However, returns, institutional trading, and sentiment are contemporaneously related in various ways. Multivariate regressions test whether institutional order flow predicts announcement day returns and sentiment after controlling

for the other market variables. Table 4 documents the predictability of news announcement returns (Panel A), news sentiment (Panel B), and institutional order flow on news days (Panel C). As in Table 3 estimates are from panel regressions with firm fixed effects. We examine volume, defined as the log of total trading volume, along with returns, sentiment, and institutional order flow. We add value-weighted industry and market returns as controls in some specifications. In addition, we split institutional order flow into its two components: institutional buys, IBuys, and sales, ISales. This decomposition helps test whether institutions are informed about both types of news, good and bad. All explanatory variables are measured on the day prior to the news announcement (day t).

Panel A of Table 4 shows that the institutional order flow imbalance and its individual components, IBuys and ISales, predict returns on news announcement days. Moreover, columns C and F indicate that only IOF has power in predicting news announcement returns. To estimate the economic significance of this result we multiply the coefficient on the IOF, 0.652, by the standard deviation of the IOF from Table 2, 0.168, to obtain that on the news days each standard deviation of IOF predicts returns of 11 basis points. In columns D–F, IBuys and ISales are each statistically significant. The positive coefficient on IBuys and negative coefficient on ISales shows that both institutional buying and selling activity predict announcement day returns, with a slightly larger negative coefficient in absolute terms. The inclusion of market and industry returns in columns C and F shows that IOF predicts the firm-specific component of returns and news.

Campbell and Thompson (2008) provide a general approach to quantifying the economic magnitude of the value investors can gain from a variable that predicts future stock returns. Campbell and Thompson (2008) assume the following process for returns, $r_{t+1} = \mu + x_t + \varepsilon_{t+1}$, where x_t is *any* observable variable. A single-period mean–variance investor with a risk aversion coefficient γ optimally chooses to invest fraction α of his wealth:

$$\alpha = \frac{1}{\gamma} \frac{\mu}{\sigma_x^2 + \sigma_\varepsilon^2} = \frac{S^2}{\mu\gamma},$$

where $S^2 \equiv \mu^2 / (\sigma_x^2 + \sigma_\varepsilon^2)$ is the unconditional Sharpe ratio of the portfolio. The investor earns an average excess return of

$$\alpha = \frac{1}{\gamma} \frac{\mu^2 + \sigma_x^2}{\sigma_\varepsilon^2} = \frac{1}{\gamma} \frac{S^2 + R^2}{1 - R^2} \quad \text{where } R^2 \equiv \frac{\sigma_x^2}{\sigma_x^2 + \sigma_\varepsilon^2}$$

is the R^2 statistic for the regression of excess return on the predictor variable x_t. The proportional increase in expected return from observing x_t is $(R^2/(1-R^2))(1+S^2)/S^2$. Utilizing the Campbell and Thompson proportional increase in expected return from the regressions in Table 4 yields how much an investor could benefit when considering investing in individual stocks knowing yesterday's institutional IOF. For column A in Panel A in Table 4 expected returns increase by 41.6% when observing IOF. In contrast, observing past returns yields an increase of 10.5% and past sentiment provides an increase of 2.8%. Hence, the statistical predictability of returns based on IOF can be economically meaningful.

Table 4

Returns, news sentiment, and institutional trading on announcement days.

   The table documents the predictability of news announcement returns (Panel A), news sentiment (Panel B), and institutional order flow (Panel C) on news announcement days. Estimates are from panel regressions with firm fixed effects. IOF denotes institutional order flow, IBuys (ISales) are institutional purchases (sales), Return is the daily stock return, Sentiment is the news sentiment, Volume is the log of total trading volume, Industry return is the value-weighted industry return (using Fama-French industry classifications) excluding the company's stock, and Market return is the value-weighted market return excluding the company's stock. All explanatory variables are measured on the day prior to the news announcement, except for industry return and market return. The sample contains all 1,667 NYSE stocks on news days from 2003 through 2005. The number of observations is 126,148. Observations are value-weighted. Standard errors are robust to heteroskedasticity and clustering and reported in parentheses. Levels of significance are denoted by * (10%), ** (5%), and *** (1%).

	(A)	(B)	(C)	(D)	(E)	(F)
Panel A: Return$_t$						
IOF$_{t-1}$	0.652***	0.693***	0.541***			
	(0.170)	(0.184)	(0.114)			
IBuys$_{t-1}$				0.632***	0.668***	0.518***
				(0.170)	(0.184)	(0.112)
ISales$_{t-1}$				−0.693***	−0.731***	−0.575***
				(0.173)	(0.186)	(0.117)
Return$_{t-1}$		−0.015	0.002		−0.015	0.002
		(0.014)	(0.007)		(0.014)	(0.007)
Sentiment$_{t-1}$		0.034	0.037*		0.034	0.037*
		(0.023)	(0.022)		(0.023)	(0.022)
Volume$_{t-1}$		−0.024	−0.004		0.001	0.018
		(0.030)	(0.010)		(0.036)	(0.013)
Industry return$_t$			0.715***			0.715***
			(0.075)			(0.075)
Market return$_t$			0.266***			0.266***
			(0.073)			(0.073)
F-statistic	54.801	16.548	2,599.178	30.684	15.003	2,231.099
Panel B: Sentiment$_t$						
IOF$_{t-1}$	0.079***	0.052**	0.052**			
	(0.022)	(0.022)	(0.022)			
IBuys$_{t-1}$				0.071***	0.046**	0.046**
				(0.022)	(0.022)	(0.022)
ISales$_{t-1}$				−0.095***	−0.061***	−0.061***
				(0.023)	(0.023)	(0.023)
Return$_{t-1}$		0.009***	0.009***		0.009***	0.008***
		(0.001)	(0.001)		(0.001)	(0.001)
Sentiment$_{t-1}$		0.167***	0.167***		0.167***	0.167***
		(0.011)	(0.011)		(0.011)	(0.011)
Volume$_{t-1}$		−0.013*	−0.012*		−0.007	−0.007
		(0.007)	(0.007)		(0.009)	(0.009)
Industry return$_t$			0.016***			0.016***
			(0.003)			(0.003)
Market return$_t$			−0.011***			−0.011***
			(0.004)			(0.004)
F-statistic	18.630	186.550	128.282	40.457	162.233	119.292
Panel C: IOF$_t$						
IOF$_{t-1}$	0.257***	0.247***	0.249***			
	(0.011)	(0.011)	(0.010)			
IBuys$_{t-1}$				0.259***	0.250***	0.252***
				(0.011)	(0.011)	(0.010)
ISales$_{t-1}$				−0.254***	−0.244***	−0.246***
				(0.011)	(0.011)	(0.011)
Return$_{t-1}$		0.004***	0.004***		0.004***	0.004***
		(0.000)	(0.000)		(0.000)	(0.000)
Sentiment$_{t-1}$		−0.001**	−0.001**		−0.001**	−0.001**
		(0.000)	(0.000)		(0.000)	(0.000)
Volume$_{t-1}$		0.002**	0.002**		−0.001	−0.001
		(0.001)	(0.001)		(0.001)	(0.001)
Industry return$_t$			0.001***			0.001***
			(0.000)			(0.000)
Market return$_t$			−0.011***			−0.011***
			(0.001)			(0.001)
F-statistic	1,056.978	521.733	742.059	546.570	424.529	642.768

Table A1 in the Appendix provides the specifications used in columns B and E in Table 4 for industries and the market. We use value-weighted aggregates for the 12 Fama-French industries and, respectively, the market. The results show support for institutions being informed, on top of firm-specific news, about industry returns and industry-wide news sentiment. However, Table A1 finds little support for institutions being informed about market-wide news. Market-wide news is further studied below in the context of macroeconomic news stories.

Panel B of Table 4 shows that IOF, IBuys, and ISales predict news sentiment. A possible alternative explanation for why IOF predicts the announcement day sentiment is that news agencies communicate with institutions, learn about institutions' trading, and interpret the direction of institutions' trading as information. Related issues could lead to sentiment being persistent as well as sentiment responding positively to past returns; i.e., higher returns in the past predict higher future sentiment. To find evidence consistent with institutions being privately informed, Section 5.3 examines earnings announcements using only the announced earnings and not sentiment.

Panel C of Table 4 shows persistence in IOF as well as its individual components. This is consistent with institutions spreading their trades over time, as predicted by Kyle (1985). It also indicates that institutions follow momentum strategies around news announcements as institutions increase purchases of past winners and sell more of past losers.

As a robustness check, in the Appendix we use the Table 4 specification on portfolios sorted based on several characteristics: (i) size, reported in Table A2; (ii) number of analysts following the firm, reported in Table A3; (iii) number of firms affected by news, reported in Table A4; (iv) Gompers, Ishii, and Metrick (GIM) governance index, reported in Table A5; and (v) Fama-French 12 industry portfolios, reported in Table A6. We find similar results in all these cases, though we lack statistical power in some of the specifications. The positive coefficients on IOF predicting returns and sentiment, respectively, are largest for firms with large market capitalization, many analysts, weak governance (high GIM index), and when many firms are mentioned in the news story. This evidence suggests that institutions are informed about news in stocks with a wide variety of characteristics and information environments.

Overall, we find that institutions trade in the right direction before a news announcement. IOF and IBuys predict positive announcement returns, positive sentiment, and more institutional buying on news days. ISales predicts negative announcement returns, negative sentiment, and more selling on news days.

4.2. Calendar-time evidence

Using the same institutional data, Boehmer and Wu (2008) and Boulatov, Hendershott, and Livdan (2013) find that institutional trading predicts returns. It is possible that institutions are informed about future price movements that are unrelated to news. To test whether IOF predicts returns more on news days we extend the return panel regressions with firm fixed effects in Panel A of Table 4 in two ways. First, we include all days in our sample period rather than just news days. Second, we interact the lagged (day $t-1$) explanatory variables with the dummy variable for news on day t. These enable the separate measurement of forecasting day t returns on news and non-news days and facilitate comparison across days.

Column A in Table 5 shows that lagged IOF predicts returns on average days. The coefficient on IOF in Table 5 is substantially smaller than the comparable coefficient in Table 4, which would arise if IOF offers additional predictability of returns on days with news. The news day variable interacted with IOF, IOF*Newsday, examines this directly. The coefficients on both IOF and IOF*Newsday are statistically significantly different from zero. Furthermore, the coefficient on IOF*Newsday tends to be larger than the coefficient on IOF, implying that the same amount of IOF predicts returns at least twice as large on news days as compared to non-news days. As with Table 4, to explore the economic significance of this difference we multiply the coefficient on the IOF, 0.259, by the standard deviation of IOF from Table 2, 0.168, to obtain that on non-news days a standard deviation of IOF predicts returns of 4.4 basis points. The coefficient on the interaction term IOF*Newsday, 0.381, multiplied by the standard deviation of the IOF is equal to 6.4 basis points. Adding the additional 6.4 news-day effect to the overall 4.4 effect shows that IOF predicts returns more than double on news days relative to days without news.

Similar to the extended specifications in Table 4, columns B–F in Table 5 incorporate additional variables into the return predictability regression, utilize Fama-French-Carhart four-factor model-adjusted returns, and separate IOF into IBuys and ISales. All columns in Table 5 show that the institutional order flow imbalance and its individual components, IBuys and ISales, predict returns on both non-news and news days and that the predictability appears greater on news days.

Consistent with prior work, e.g., Tetlock (2007), columns B–F in Table 5 also show that sentiment positively predicts returns on average days. In contrast, the coefficient on sentiment in Table 4 is positive, but not statistically significant. The coefficient on sentiment*Newsday is negative, but not statistically significant. The sum of the coefficients on sentiment and sentiment*Newsday is positive, but not statistically significant, consistent with Table 4. Thus, we find evidence that sentiment predicts returns on non-news days, but not on news days.

Up to this point all regression analysis examines the relation of news and institutional trading across consecutive days. Dynamics at longer lags or the lagged and contemporaneous relations between news sentiment and institutional trading could lead to our findings thus far without institutions being informed about the news. For example, if sentiment predicts IOF the next day and then returns the day after that, our specifications using only one lag would find an association between returns and lagged IOF, which may not be associated with institutions being informed. Similarly, sentiment earlier in the trading day could lead to IOF later that same day through institutions interpreting news and trading on their posterior estimate of value. Another possibility is that institutional trading is

Table 5

Returns, news sentiment, and institutional trading.

The table documents the predictability of stock returns. Estimates are from panel regressions with firm fixed effects. IOF denotes institutional order flow, IBuys (ISales) are institutional purchases (sales), Return is the daily stock return, $Return^{FF4}$ is the Fama-French-Carhart four-factor model-adjusted return, Sentiment is the news sentiment, and Volume is the log of total trading volume. Newsday is the contemporaneous news announcement indicator variable. All explanatory variables are lagged by one day, except for Newsday. The sample contains all 1,667 NYSE stocks on all 755 trading days from 2003 through 2005. The number of observations is 1,096,514. Observations are value-weighted. Standard errors are robust to heteroskedasticity and clustering and reported in parentheses. Levels of significance are denoted by * (10%), ** (5%), and *** (1%).

	$Return_t$ (A)	$Return_t$ (B)	$Return_t^{FF4}$ (C)	$Return_t$ (D)	$Return_t$ (E)	$Return_t^{FF4}$ (F)
IOF_{t-1}	0.259***	0.288***	0.222***			
	(0.050)	(0.061)	(0.025)			
$IOF_{t-1}*Newsday_t$	0.381***	0.398***	0.209**			
	(0.144)	(0.153)	(0.090)			
$IBuys_{t-1}$				0.254***	0.288***	0.218***
				(0.051)	(0.062)	(0.025)
$IBuys_{t-1}*Newsday_t$				0.376***	0.375**	0.195**
				(0.144)	(0.152)	(0.090)
$ISales_{t-1}$				−0.268***	−0.288***	−0.228***
				(0.050)	(0.061)	(0.026)
$ISales_{t-1}*Newsday_t$				−0.393***	−0.431***	−0.228**
				(0.146)	(0.154)	(0.093)
$Return_{t-1}$		−0.019	−0.011***		−0.019	−0.011***
		(0.013)	(0.003)		(0.013)	(0.003)
$Return_{t-1}*Newsday_t$		0.005	0.020***		0.005	0.020***
		(0.008)	(0.007)		(0.008)	(0.007)
$Sentiment_{t-1}$		0.079***	0.061***		0.080***	0.061***
		(0.022)	(0.016)		(0.022)	(0.016)
$Sentiment_{t-1}*Newsday_t$		−0.046	−0.022		−0.046	−0.022
		(0.032)	(0.026)		(0.032)	(0.026)
$Volume_{t-1}$		−0.012	0.018**		−0.012	0.022**
		(0.026)	(0.008)		(0.031)	(0.011)
$Volume_{t-1}*Newsday_t$		0.008	−0.007		0.032	0.007
		(0.015)	(0.011)		(0.021)	(0.015)
$Newsday_t$	0.015	0.015	0.021**	0.021	0.035	0.033***
	(0.021)	(0.020)	(0.009)	(0.022)	(0.022)	(0.011)
F-statistic	58.989	28.845	24.010	36.246	24.158	19.917

positively autocorrelated, has a positive price impact, and causes subsequent news. In this case institutional trading may not predict returns beyond predicting future institutional trading's impact on price.

To test the robustness of our interpretation of the regression results in Tables 4 and 5, we next explicitly account for the contemporaneous and lagged relations among returns, news sentiment, and institutional trading using panel vector autoregressions (VARs). By using additional lags the VAR can control for more complex dynamic relations among the variables. Examining the impulse response functions associated with the VAR enables us to test for the responses of the variables to innovations in the other variables. The impulse responses can also be calculated while eliminating/orthogonalizing contemporaneous innovations in the variables. These allow testing of the impact of an innovation in IOF under the most conservative assumptions.

For each firm i and time t we combine three variables into a 3×1 vector $\mathbf{y}_{it}$, $\mathbf{y}_{it} = (IOF_{it}, Return_{it}, Sentiment_{it})'$. Following Holtz-Eakin, Newey, and Rosen (1988), we allow the individual components of $\mathbf{y}_{it}$ to be autocorrelated and jointly endogenously determined by specifying

$$\mathbf{y}_{it} = \boldsymbol{\alpha}_i + \sum_{l=1}^{L} \boldsymbol{\lambda}^l \mathbf{y}_{it-l} + \boldsymbol{\varepsilon}_{it}, \tag{6}$$

where $\boldsymbol{\alpha}_i$ is a 3×1 vector of firm-specific intercepts, $\boldsymbol{\lambda}^l$, $l = 1, \ldots, L$, are 3×3 coefficient matrices, and $\boldsymbol{\varepsilon}_{it}$ is a 3×1 vector of innovations.

The main advantage of the specification (6) is that it allows us to relax the "pooling" constraint that the time series relationship of IOF_{it}, $Return_{it}$, and $Sentiment_{it}$ is the same for each firm. One way to relax this constraint is to allow for an intercept $\boldsymbol{\alpha}_i$ in Eq. (6) that varies across firms. Changes in the intercept of a stationary VAR correspond to changes in the means of IOF_{it}, $Return_{it}$, and $Sentiment_{it}$. We do not use a time-varying intercept, as making the intercept time dependent is most useful when the time series is short. We have 756 observations per firm. Furthermore, we keep the stationarity constraint in Eq. (6) because all three components of the vector $\mathbf{y}$ are stationary. In addition, we allow for individual heterogeneity by making the variance of the innovations in Eq. (6) heteroskedastic across firms. Changes in the innovation variance of a VAR translate into changes in the variance of the variables, so that allowing for cross-sectional heteroskedasticity in the innovation variance allows for individual heterogeneity in the variability of IOF_{it}, $Return_{it}$, and $Sentiment_{it}$.

The firm fixed effects $\boldsymbol{\alpha}_i$ in Eq. (6) are correlated with the regressors due to lags of the dependent variables, leading to biased ordinary least squares (OLS) estimates. However, the specification of Eq. (6) as a projection equation implies that the error term $\boldsymbol{\varepsilon}_{it}$ satisfies the orthogonality conditions

$E[\mathbf{y}_{is}\boldsymbol{\varepsilon}_{it}] = E[\boldsymbol{\alpha}_i\boldsymbol{\varepsilon}_{it}] = 0$ for $s < t$. These orthogonality conditions imply that lagged values of $\mathbf{y}$ qualify as instrumental variables for Eq. (6). In order to use the orthogonality conditions to identify the parameters of Eq. (6), we must deal with the presence of the unobserved individual effect, $\boldsymbol{\alpha}_i$. It is well known that in models with lagged dependent variables it is inappropriate to treat individual fixed effects as constants to be estimated. Therefore, we apply the forward orthogonal deviations transformation, or Helmert transform, as in Arellano and Bover (1995), in Eq. (6) to eliminate individual fixed effects. This transformation preserves the orthogonality between transformed error terms and lagged regressors, so we can continue to use lagged regressors as instruments.

Let
$$\mathbf{y}_{it}^* = ((T-t)/(T-t+1))^{1/2}(\mathbf{y}_{it} - (1/(T-t))\textstyle\sum_{s=t+1}^{T}\mathbf{y}_{is})$$
be the forward orthogonal deviation of $\mathbf{y}_{it}$. Then (6) reduces to

$$\mathbf{y}_{it}^* = \sum_{l=1}^{L}\lambda^l\mathbf{y}_{it-l}^* + \boldsymbol{\varepsilon}_{it}^*, \tag{7}$$

where $\boldsymbol{\varepsilon}_{it}^* = ((T-t)/(T-t+1))^{1/2}(\boldsymbol{\varepsilon}_t - (1/(T-t))\sum_{s=t+1}^{T}\boldsymbol{\varepsilon}_s)$ is the transformed error term. It immediately follows that moment conditions $E[\mathbf{y}_{is}\boldsymbol{\varepsilon}_{it}] = 0$ imply $E[\mathbf{y}_{is}\boldsymbol{\varepsilon}_{it}^*] = 0$ for $s < t$.

We estimate the coefficients in (7) by the system generalized method of moments (GMM) with the lagged untransformed variables $\mathbf{y}_{is}$ for $s < t$ available as instruments (Holtz-Eakin, Newey, and Rosen, 1988). As in Hasbrouck (1991), we do not set the lag length optimally using the Akaike or Schwarz information criteria. Instead we choose $L=3$ lags for all stocks because this lag structure is sufficient to eliminate all the serial correlation in the data (see Panel B). Table A7 in the Appendix shows that the results that follow do not rely on this particular lag structure. For the just-identified system, the vector of instrumental variables (IV) to identify the parameters λ^l of Eq. (7) is $\mathbf{Z}_{it} = [\mathbf{1}, \mathbf{y}_{it-1}, ..., \mathbf{y}_{it-L}]$. With this set of instruments, the system GMM estimator is equivalent to equation-by-equation IV estimation on the forward demeaned $\mathbf{y}_{it}^*$ with instruments $\mathbf{Z}_{it}$.

Table 6 reports estimates of specification (7). The results generally are consistent with the evidence from Tables 4 and 5. Consistent with the event-study graphs in Fig. 3 institutional order flow predicts returns at more than one lag. Lag one returns negatively predict returns while lag one sentiment positively predicts returns. Regression coefficients on one-day lagged returns and sentiment have the same signs as in column B, Panel A of Table 4 but in Table 6 they are both statistically significant. These differences could arise from the VARs using both news and non-news days. The lag one return coefficient in Table 6 of -0.005 is smaller than the corresponding -0.015 coefficient in Table 4, but the VAR coefficient may be more precisely estimated due to the almost ten times larger sample size when non-news days are included. In Table 5 returns have a negative coefficient, but are not statistically significant. The VAR may improve the efficiency of the estimation by better modeling dynamics between the variables. Both returns and sentiment lose their predictive power of future returns at a two-day horizon.

Table 6
Vector autoregressions of returns, news sentiment, and institutional trading.

The table reports estimates from panel vector autoregressions (6) with firm fixed effects. The estimates are obtained using system GMM estimation as described in Section 4.2. The dependent variables are institutional order flow (IOF), stock returns, and news sentiment. We set the lag length $L=3$ that minimizes the Bayesian information criterion (BIC) and Akaike information criterion (AIC). The VAR diagnostics in Panel B report the BIC, AIC, mean-squared error (MSE), and residual auto- and cross-correlations. The sample contains all 1,667 NYSE stocks on all 755 trading days from 2003 through 2005. The number of observations is 1,096,514. Standard errors are reported in parentheses. Levels of significance are denoted by * (10%), ** (5%), and *** (1%).

	IOF_t	$Return_t$	$Sentiment_t$
Panel A: Estimates			
IOF_{t-1}	0.236***	0.171***	0.002**
	(0.008)	(0.017)	(0.001)
IOF_{t-2}	0.065***	0.041**	−0.001
	(0.006)	(0.016)	(0.001)
IOF_{t-3}	0.055***	−0.022	0.002**
	(0.005)	(0.015)	(0.001)
$Return_{t-1}$	0.003***	−0.005***	0.001***
	(0.000)	(0.001)	(0.000)
$Return_{t-2}$	0.000***	−0.002	0.000
	(0.000)	(0.001)	(0.000)
$Return_{t-3}$	−0.001***	−0.004***	0.000
	(0.000)	(0.001)	(0.000)
$Sentiment_{t-1}$	0.001	0.093***	0.100***
	(0.001)	(0.014)	(0.002)
$Sentiment_{t-2}$	−0.001	0.020	0.028***
	(0.001)	(0.013)	(0.002)
$Sentiment_{t-3}$	−0.001	0.005	0.033***
	(0.001)	(0.013)	(0.002)
Panel B: Diagnostics			
AIC/BIC		−6.264/ −6.264	
MSE	0.161	1.943	0.140
Residual auto-correlation	0.001	−0.001	0.007***
Residual cross-correlation:			
IOF_t		0.056***	0.001
$Return_t$			0.050***

As in Table 4 institutional order flow is persistent. Positive returns today and yesterday both predict higher institutional order imbalances the following day. Sentiment today negatively predicts higher institutional order flow, as can be seen in Panel C of Table 4. This effect is not present in Table 6. As before both lagged returns and lagged sentiment predict sentiment.

We apply a Cholesky decomposition to the residual variance-covariance matrix in order to obtain orthogonalized impulse responses. We vary the ordering of the variables, assigning contemporaneous correlation between the residuals to the variables that come earlier in the ordering. We then calculate standard errors for the impulse-response functions by Monte Carlo simulations, accounting for estimation error in the coefficients (Hamilton, 1994, Chapter 11.7). Specifically, we randomly draw a set of coefficients from a normal distribution centered around the point estimates for λ and with the same error variance–covariance matrix. For each draw we recalculate the impulse responses. We repeat this procedure 1,000 times to obtain 5% confidence bounds on the impulse responses.
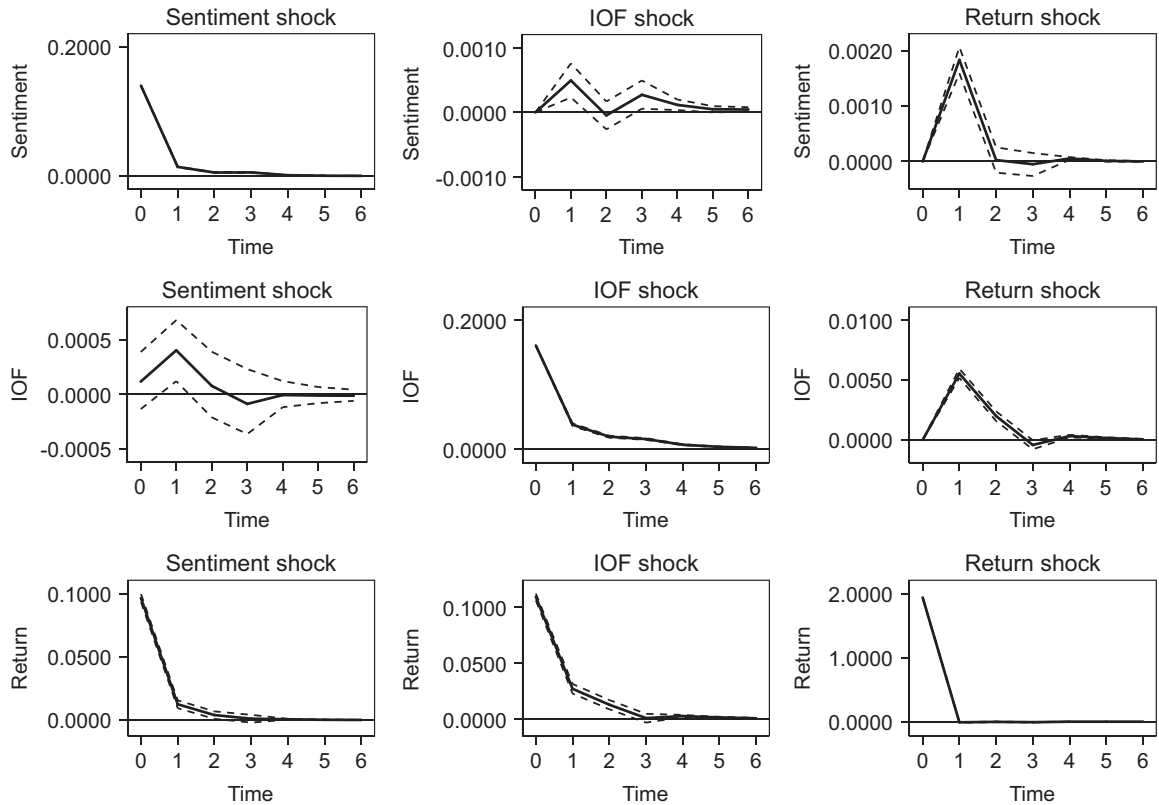
Fig. 4. Impulse responses. The figure reports the impulse response functions corresponding to the panel VAR (6) with lag length $L=3$ reported in Table 6. The estimates in Table 6 are obtained using GMM estimation as described in Section 4.2. The dependent variables are ordered in the following sequence: sentiment, IOF, return. Impulse responses correspond to a one standard deviation shock. Error bands at 5% level for the impulse responses (dashed lines) are generated using Monte-Carlo simulations with 1,000 draws.

Fig. 4 reports the impulse response functions (IRF) corresponding to the panel VAR estimated in Table 6. The dependent variables are ordered in the following sequence: sentiment, IOF, returns. The ordering assumption of sentiment, IOF, and returns attributes predictability arising from a contemporaneously correlated component between sentiment and IOF to sentiment and not IOF. Fig. A1 in the Appendix contains all possible orderings for the variables showing impulse response functions consistent with Fig. 4. The IRFs are qualitatively consistent with the coefficient estimates in Table 6 as IOF shocks lead to news sentiment and returns. IOF shocks predict future returns up to three to five subsequent days and sentiment the next day. Overall, the VAR shows that allowing for additional dynamics and longer lags continues to support the finding that IOF predicts future news sentiment and returns.

One restriction used in specification (6) is that coefficients on lags of all endogenous variables are assumed to be constant across firms; i.e., λ^l does not depend on firm index i. This restriction is referred to as static heterogeneity. The case when the lags of all endogenous variables of all firms enter the model for firm i is called dynamic heterogeneity. In this case, the estimator (7) will give inconsistent estimates of the parameters, since the error term is also likely to be correlated with the regressors. In this case, a random coefficient model is known to help to improve the quality of the estimates of coefficients. However, with dynamic

heterogeneity the GMM strategy used above is difficult to employ since it is hard to find instruments which are simultaneously correlated with the regressors and uncorrelated with the error term. There exists no general estimation procedure for the random coefficients model in the context of panel VAR and, creating an additional complication, the number of random effect variances and correlations to be estimated grows exponentially. The available estimators, which are scarce (see Canova, 2005; Canova and Pappa, 2007; Auerbach and Gorodnichenko, 2011; and more recently Calza, Monacelli, and Stracca, 2013), rely heavily on the specifics of the underlying economic model.

To keep our specification tractable, we consider the following random intercept and random slopes model. For each firm i and time t, the 3×1 vector $\mathbf{y}_{it} = (\text{IOF}_{it}, \text{Return}_{it}, \text{Sentiment}_{it})'$ is heterogeneously autocorrelated and endogenously determined as follows:

$$\mathbf{y}_{it} = \boldsymbol{\alpha}_{it} + \sum_{l=1}^{L} \boldsymbol{\lambda}_{it}^l \mathbf{y}_{it-l} + \boldsymbol{\varepsilon}_{it}, \qquad (8)$$

where $\boldsymbol{\varepsilon}_{it} \sim \mathcal{N}(0, \sigma_\varepsilon^2)$ is a 3×1 vector of innovations. In contrast to (6), the intercept $\boldsymbol{\alpha}_{it} = \boldsymbol{\alpha}_0 + \boldsymbol{\beta}\boldsymbol{X}_{it} + \boldsymbol{\epsilon}_i$ is now a random 3×1 vector that varies across firms and time linearly with the explanatory variables $\boldsymbol{X}_{it}$ and the random effects $\boldsymbol{\epsilon}_i \sim \mathcal{N}(0, \sigma_\alpha^2)$. Similarly, $\boldsymbol{\lambda}_i^l = \boldsymbol{\lambda}_0^l + \boldsymbol{\gamma}^l \boldsymbol{X}_{it} + \boldsymbol{\eta}_i^l$, $l = 1, ..., L$, are firm-specific 3×3 random coefficient matrices with $\boldsymbol{\eta}_i^l \sim \mathcal{N}(0, \sigma_{\lambda^l}^2)$. The model lends itself to maximum likelihood

estimation with the expectation–maximization (EM) algorithm (see Dempster, Laird, and Rubin, 1977). We set the lag length $L=1$ in the reported specifications, as longer lags do not change the results.

As a robustness check, we estimate our panel VAR model (6) on subsamples, i.e., we sort firms into terciles based on such firm-specific characteristics as size, governance, and number of following analysts and then compare GMM-estimated panel VAR coefficients across the terciles. As a final robustness check, we do pooling across the 12 Fama-French industries.

All our results are reported in the Appendix and we provide only a brief summary of them here. The parameter estimates for model (8) in Table A8 show there is indeed sizeable (observable and unobserved) heterogeneity across firms in the dynamic relation between IOF, stock returns, and news sentiment. All σ_{λ^l} estimates for the random effects are significantly positive. IOF consistently predicts returns and sentiment (see λ_0^l in specification A), and the coefficients are larger for larger firms that tend to have more institutional ownership (see γ^l in specification B). The coefficients on IOF are also larger in more weakly governed firms (a high GIM index represents weak shareholder power). This suggests that institutions have stronger monitoring incentives in firms with more managerial entrenchment. Alternatively, the leakage of privileged information to institutional investors is more pronounced in these types of firms. Further, IOF and returns tend to predict future returns more for higher market-to-book firms, in which opaque intangibles are more important. Finally, the return predictability from news sentiment is much smaller when the number of analysts following the company is larger.

For industry-sorted portfolios, as well as portfolios sorted on the number of analysts, we cannot reject the null in Tables A9 and A10 that all statistically significant regression coefficients on the lags are the same across different group categories. In the case of size-sorted portfolios, we cannot reject the null for all statistically significant regression coefficients on the lags except for the coefficient on the first lag of IOF in the returns regression. Consistent with the corresponding γ^l estimate, this coefficient is monotonically increasing across size portfolios. Overall, our results are robust to the above types of dynamic heterogeneity across firms and time.

5. What types of news are institutions informed on?

Up to this point to examine whether institutions are informed about news in the broadest context, the analysis has grouped all news announcements together. The results showing institutional order flow predicting sentiment and returns in Fig. 3 and Tables 4–6 are consistent with institutions having private information about the news. In this section we look at specific news categories. Instead of looking at the individual news categories as classified by Reuters we group news into broader news groups based on their content. First, motivated by the fact that asset prices behave very differently on days when important macroeconomic news is scheduled for announcement relative to other trading days (Savor and Wilson, 2014),

we investigate institutional trading around macroeconomic news. It is possible that some news announcements contain information impacting prices that is unrelated to fundamentals. Therefore, we next turn our attention to two types of news directly related to fundamentals: crisis and earnings announcements. Focusing on these news types allows us to study whether institutions are informed about the timing of news (crises have strong negative sentiment but uncertain timing), news content (earnings announcements are scheduled in advance but have uncertain content), or both. Finally, we explore the question of whether institutions trade around news stories with less relation to longer term information, which we refer to as institutions believing the "hype."

5.1. Macroeconomic news

Days when macroeconomic news is announced are special for asset pricing. Savor and Wilson (2014) demonstrate that on days when news about inflation, unemployment, or Federal Open Markets Committee interest-rate decisions is scheduled to be announced, stock market beta is strongly related to returns and a robustly positive risk-return trade-off exists. In this section we investigate how institutions trade around macroeconomic news announcements.

We define MACRO news days as days with news on one of the following five topics: Economic Indicators (ECI, mean sentiment -0.127), Federal Reserve Board (FED, mean sentiment -0.197), Government/Sovereign Debt (GVD, mean sentiment -0.103), Macroeconomics (MCE, average sentiment -0.110), and Washington/U.S. Government News (WASH, mean sentiment -0.232). MACRO news have mean sentiment of -0.189, median sentiment of -0.271, and the standard deviation of their sentiment is equal to 0.441.

If institutions are informed about macroeconomic news, institutional trading and specifically institutional selling should have predictive power for returns and sentiment around MACRO news days. We analyze the relations among IOF, news sentiment, and returns on news days for the MACRO news category as well as its five constituent news categories using specification E in Table 4.

Table 7 reports results for the MACRO news category as well as for each individual news category included in MACRO news. The reported estimates are from panel regressions with firm fixed effects and volume, defined as the log of total trading volume, as control in addition to returns, sentiment, and IBuys and ISales. All explanatory variables are measured on the day prior to the news announcement. The estimates show that institutional trading predicts returns (Panel A) on MACRO news days. For the individual news categories the return predictability is statistically significant only for the WASH category, which is by far the largest news category included in MACRO news (13,075 observations; the next largest category, GVD, has only 3,682 observations). Institutional order flow does not predict sentiment (Panel B) for the MACRO news category as a whole. While the regression coefficients on IBuys and ISales have the correct signs, they are not statistically significant. The regression coefficients on IBuys and ISales for the individual news categories are all

Table 7

Returns, news sentiment, and institutional trading on macro news days.

The table documents the predictability of news announcement day returns (Panel A), news sentiment (Panel B), and institutional order flow (Panel C) on MACRO news days. Macro days are defined as days with news on one of the following topics: ECI, FED, GVD, MCE, WASH. Table 1 provides topic definitions. Estimates are from panel regressions with firm fixed effects. IOF denotes institutional order flow, IBuys (ISales) are institutional purchases (sales), Sentiment is the news sentiment, and Volume is the log of total trading volume. All explanatory variables are measured on the day prior to the news announcement. The sample contains all 1,667 NYSE stocks on macroeconomic news days from 2003 through 2005. The number of observations in each portfolio is reported in the third row of the table. Observations are value-weighted. Standard errors are robust to heteroskedasticity and clustering and reported in parentheses. Levels of significance are denoted by * (10%), ** (5%), and *** (1%).

		MACRO categories				
N	MACRO 17,586	ECI 1,812	FED 990	GVD 3,682	MCE 1,918	WASH 13,075
Panel A: Return$_t$						
IBuys$_{t-1}$	0.756**	0.518	0.872	0.736	−0.028	1.069**
	(0.359)	(0.803)	(1.257)	(0.581)	(1.099)	(0.456)
ISales$_{t-1}$	−0.810**	−0.787	−0.883	−0.999	−0.406	−1.084**
	(0.374)	(0.801)	(1.311)	(0.687)	(1.041)	(0.470)
Return$_{t-1}$	0.018	0.017	−0.052	0.028	0.022	−0.005
	(0.020)	(0.040)	(0.042)	(0.036)	(0.050)	(0.025)
Sentiment$_{t-1}$	0.056	0.048	0.221	0.273***	0.154	−0.031
	(0.055)	(0.186)	(0.288)	(0.097)	(0.208)	(0.069)
Volume$_{t-1}$	−0.061	0.053	0.038	0.104	0.257	−0.103
	(0.050)	(0.164)	(0.205)	(0.132)	(0.162)	(0.063)
F-statistic	2.424	0.331	0.401	1.917	0.766	2.020
Panel B: Sentiment$_t$						
IBuys$_{t-1}$	0.121	0.481**	−0.224	0.066	−0.061	0.174
	(0.094)	(0.208)	(0.495)	(0.164)	(0.154)	(0.133)
ISales$_{t-1}$	−0.122	−0.467**	0.221	−0.005	0.047	−0.192
	(0.096)	(0.194)	(0.528)	(0.151)	(0.160)	(0.139)
Return$_{t-1}$	0.011***	0.022	0.040***	0.023***	0.033**	−0.001
	(0.004)	(0.017)	(0.015)	(0.006)	(0.016)	(0.006)
Sentiment$_{t-1}$	0.173***	0.189***	0.256***	0.234***	0.211***	0.135***
	(0.019)	(0.057)	(0.098)	(0.043)	(0.075)	(0.019)
Volume$_{t-1}$	−0.035***	−0.102*	0.004	−0.115***	−0.011	−0.013
	(0.011)	(0.053)	(0.070)	(0.038)	(0.045)	(0.015)
F-statistic	25.214	6.696	4.520	19.570	3.814	9.877
Panel C: IOF$_t$						
IBuys$_{t-1}$	0.210***	0.140**	0.212***	0.191***	0.159***	0.239***
	(0.021)	(0.055)	(0.076)	(0.031)	(0.056)	(0.020)
ISales$_{t-1}$	−0.201***	−0.149***	−0.217***	−0.167***	−0.157***	−0.238***
	(0.023)	(0.050)	(0.070)	(0.042)	(0.053)	(0.019)
Return$_{t-1}$	0.004***	0.003***	0.004***	0.003***	0.003***	0.004***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
Sentiment$_{t-1}$	−0.002**	−0.006	−0.004	−0.003	−0.001	−0.003**
	(0.001)	(0.004)	(0.007)	(0.003)	(0.003)	(0.001)
Volume$_{t-1}$	0.001	0.005	0.003	−0.006	0.004	0.002
	(0.002)	(0.005)	(0.007)	(0.004)	(0.003)	(0.002)
F-statistic	58.226	5.904	3.872	18.615	9.396	46.282

over the place and are not statistically significant with the exception of ECI news. The institutional order flow predicts the sentiment of the ECI news: ISales predicts a negative surprise whereas IBuys predicts a positive surprise, and both coefficients are statistically significant.

Overall, while institutional trading predicts returns on MACRO news days, institutions do not trade in the direction of macro news with the exception of economic indicators.

5.2. Crisis surprises

We define "crisis" as an unexpected value-destroying event. A broad definition allows us to concentrate on specific news categories instead of specific individual news events,

thus increasing the number of observations and representativeness. We select news categories for the CRISIS category based on two criteria: (1) content has to be about unexpected value-destroying events; (2) both average and median sentiment for the entire news category has to be negative. There are many news categories that fit one of the two selection criteria but only eight news categories fit both: (1) bankruptcies (BKRT, mean sentiment equal to −0.353), which contains news on corporate insolvency, bankruptcy filings, and court rulings; (2) management issues/policy (MNGISS, mean sentiment equal to −0.161), which contains news on executive pay, bonuses, governance, and accounting irregularities; (3) jobs (JOB, mean sentiment equal to −0.153), which contains news on (un)employment,

Table 8

Returns, news sentiment, and institutional trading on crisis days.

The table documents the predictability of news announcement day returns (Panel A), news sentiment (Panel B), and institutional order flow (Panel C) on crisis days. Crisis days are defined as days with news on one of the following topics: BKRT, MNGISS, JOB, CRIM, JUDIC, REGS, CDV, WEA. Table 1 provides topic definitions. Estimates are from panel regressions with firm fixed effects. IOF denotes institutional order flow, IBuys (ISales) are institutional purchases (sales), Sentiment is the news sentiment, and Volume is the log of total trading volume. All explanatory variables are measured on the day prior to the news announcement. The sample contains all 1,667 NYSE stocks on crisis news days from 2003 through 2005. The number of observations in each portfolio is reported in the third row of the table. Observations are value-weighted. Standard errors are robust to heteroskedasticity and clustering and reported in parentheses. Levels of significance are denoted by * (10%), ** (5%), and *** (1%).

		CRISIS categories							
N	CRISIS 31,633	BKRT 2,069	MNGISS 7,338	JOB 8,766	CRIM 3,925	JUDIC 3,407	REGS 14,270	CDV 2,466	WEA 2,234
Panel A: Return$_t$									
IBuys$_{t-1}$	0.738***	1.350	0.455	0.788	1.278*	1.047	1.127**	−0.162	−0.190
	(0.277)	(0.961)	(0.593)	(0.522)	(0.701)	(0.964)	(0.453)	(0.739)	(0.432)
ISales$_{t-1}$	−0.872***	−1.652	−0.621	−0.978*	−1.444**	−1.303	−1.287***	−0.265	0.226
	(0.289)	(1.160)	(0.575)	(0.539)	(0.709)	(0.925)	(0.488)	(0.759)	(0.461)
Return$_{t-1}$	−0.003	0.069	−0.007	−0.008	−0.029	−0.025	0.009	0.070***	−0.095**
	(0.017)	(0.053)	(0.019)	(0.026)	(0.045)	(0.034)	(0.019)	(0.025)	(0.043)
Sentiment$_{t-1}$	0.047	0.107	0.062	0.106	−0.027	0.036	0.013	−0.478***	0.086
	(0.038)	(0.186)	(0.115)	(0.121)	(0.140)	(0.116)	(0.053)	(0.148)	(0.137)
Volume$_{t-1}$	−0.012	−0.013	0.016	−0.000	0.092	0.059	−0.000	0.224	−0.087
	(0.046)	(0.121)	(0.088)	(0.084)	(0.123)	(0.162)	(0.058)	(0.165)	(0.100)
F-statistic	4.970	0.661	0.875	2.615	1.103	1.034	2.374	3.516	1.128
Panel B: Sentiment$_t$									
IBuys$_{t-1}$	0.156**	0.718**	0.224	0.118	0.038	0.236	0.094	0.058	0.382
	(0.065)	(0.327)	(0.142)	(0.110)	(0.183)	(0.162)	(0.086)	(0.246)	(0.298)
ISales$_{t-1}$	−0.171***	−0.807**	−0.226	−0.117	−0.019	−0.226	−0.107	−0.072	−0.383
	(0.062)	(0.355)	(0.150)	(0.111)	(0.182)	(0.155)	(0.086)	(0.266)	(0.287)
Return$_{t-1}$	0.003	−0.010	0.005	−0.001	−0.002	0.001	0.006	0.008	−0.006
	(0.003)	(0.012)	(0.007)	(0.006)	(0.006)	(0.008)	(0.004)	(0.010)	(0.013)
Sentiment$_{t-1}$	0.151***	0.095**	0.185***	0.139***	0.131***	0.009	0.128***	0.221***	0.197***
	(0.020)	(0.046)	(0.056)	(0.022)	(0.040)	(0.032)	(0.022)	(0.035)	(0.056)
Volume$_{t-1}$	−0.005	0.056	−0.035	−0.008	0.012	−0.024	0.000	−0.018	0.050
	(0.014)	(0.039)	(0.025)	(0.036)	(0.021)	(0.022)	(0.013)	(0.047)	(0.033)
F-statistic	26.929	2.014	9.552	3.632	3.566	0.847	10.997	3.420	3.061
Panel C: IOF$_t$									
IBuys$_{t-1}$	0.241***	0.192***	0.256***	0.206***	0.210***	0.248***	0.243***	0.191***	0.206**
	(0.016)	(0.036)	(0.022)	(0.026)	(0.022)	(0.036)	(0.020)	(0.049)	(0.088)
ISales$_{t-1}$	−0.237***	−0.206***	−0.257***	−0.207***	−0.203***	−0.248***	−0.243***	−0.179***	−0.198**
	(0.016)	(0.038)	(0.023)	(0.027)	(0.023)	(0.037)	(0.019)	(0.048)	(0.085)
Return$_{t-1}$	0.004***	0.004***	0.004***	0.004***	0.004***	0.004***	0.004***	0.003***	0.005***
	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)
Sentiment$_{t-1}$	−0.001*	0.004*	−0.002	−0.000	−0.001	−0.002	−0.003***	−0.008*	0.002
	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.003)	(0.001)	(0.004)	(0.005)
Volume$_{t-1}$	0.001	0.007	0.001	0.001	0.001	0.005	0.002	0.002	0.003
	(0.002)	(0.005)	(0.003)	(0.003)	(0.004)	(0.004)	(0.002)	(0.007)	(0.005)
F-statistic	119.742	16.112	31.455	32.187	17.114	16.659	57.309	9.657	7.145

labor disputes, strikes, and unions; i.e., issues related to increases in input costs and interruptions in production and(or) deliveries of products; (4) crime (CRIM, mean sentiment equal to −0.460), which contains news on civil and criminal lawsuits, corporate crime, fraud, murder, and criminals; (5) judicial processes/court cases (JUDIC, mean sentiment equal to −0.459) which contains news on court decisions; (6) news on regulatory issues (REGS, mean sentiment equal to −0.261); (7) credit default swaps news (CDV, mean sentiment equal to −0.195); i.e., news affecting market perception about companies' solvency; (8) weather reports, warnings, forecasts, and statistics (WEA, mean sentiment equal to −0.262); i.e., news on natural disasters affecting costs of inputs, deliveries, and production as well as leading

to property destruction. The CRISIS news category formed by pooling these eight news categories has mean sentiment of −0.221 and median sentiment of −0.335 with standard deviation of 0.436. The Martha Stewart example in the introduction is part of the CRISIS category.

If institutions are informed about crises, institutional trading and specifically institutional selling should have predictive power for returns and sentiment around crisis days. In the extreme case of crisis the sentiment of all news is equal to −1 with certainty. In this case only return and institutional trading predictability around the event days should be considered. In our sample there is enough variation in the crisis sentiment on news days to also report sentiment predictability

results. We follow the previous section and use specification E in Table 4 as a blueprint for our analysis.

Table 8 reports estimates from panel regressions with firm fixed effects and volume, defined as the log of total trading volume, returns, sentiment, and IBuys and ISales. All explanatory variables are measured on the day prior to the news announcement. The estimates show that institutional trading predicts returns (Panel A) and sentiment (Panel B) for the CRISIS news category. Institutions trade in the right direction before a crisis announcement. ISales predicts a negative surprise whereas IBuys predicts a positive surprise. For the individual news categories the return predictability is statistically significant for the CRIM and REGS category and marginally significant for the JOB category. Institutional order flow predicts sentiment only for the BKRT news category (coefficients on IBuys and ISales are statistically significant and have correct signs). However, the regression coefficients on IBuys and ISales for the individual news categories all have correct signs; i.e., positive sign for IBuys and negative sign for ISales. Another interesting result is that unlike Tetlock (2007) and Tetlock, Saar-Tsechansky, and Macskassy (2008), news sentiment does not predict individual returns on crisis days, at least not statistically significantly so.

Overall, our results support the hypothesis that institutions are informed about crisis. Because our sample period is calm for the market as a whole, the crisis events provide evidence on institutions' role in firm governance during times of stress.

5.3. Earnings announcement surprises

This section provides further evidence that institutions have private value-relevant information by studying firms' earnings announcements in more detail. If institutions are informed about corporate performance, institutional trading should have predictive power for analyst forecast errors around earnings days. The standardized unanticipated earnings, SUE, score is a commonly used measure to quantify the surprise in the marketplace. The SUE score measures the deviation of the announced earnings from the mean analyst estimate. We compute the SUE score by aggregating the published earnings forecasts from the Thomson Reuters Institutional Brokers' Estimate System (I/B/E/S).

The SUE score measures the number of standard deviations the actual reported earnings differ from the I/B/E/S mean estimates for a company for the current fiscal period. The SUE score for stock i on the announcement day is calculated as

$$SUE_i = \frac{ER_i - E[\widehat{ER}_i]}{\sigma(\widehat{ER}_i)}, \tag{9}$$

where the surprise mean, $E[\widehat{ER}_i]$, is the arithmetic average of analysts' estimates on the release date of the quarterly earnings, ER_i. The surprise standard deviation, $\sigma(\widehat{ER}_i)$, measures the dispersion in analysts' estimates at the time of the earnings announcement by the standard deviation

Table 9
Returns, institutional trading, and earnings surprises.

The table documents the predictability of earnings surprises by institutional order flow. Estimates are from panel regressions with firm fixed effects. Earnings surprises are measured by the standardized unanticipated earnings, SUE, score. The SUE score is calculated as follows:

$$SUE_i = \frac{ER_i - E[\widehat{ER}_i]}{\sigma(\widehat{ER}_i)},$$

where the surprise mean, $E[\widehat{ER}_i]$, is the arithmetic average of analysts' estimates on the release date of the quarterly earnings, ER_i, and $\sigma(\widehat{ER}_i)$ is the standard deviation of individual analyst estimates about the average estimate. IOF denotes institutional order flow, IBuys (ISales) are institutional purchases (sales), Sentiment is the news sentiment, and Volume is the log of total trading volume. All explanatory variables are measured on the day prior to the news announcement. The sample contains all 1,667 NYSE stocks on earnings announcement days from 2003 through 2005. The number of observations is 9,144. Observations are value-weighted. Standard errors are robust to heteroskedasticity and clustering and reported in parentheses. Levels of significance are denoted by * (10%), ** (5%), and *** (1%).

Panel A: Correlations

	$Return_t$	$Sentiment_t$	IOF_t	$IBuys_t$	$ISales_t$	$Volume_t$
SUE_t	0.3390***	0.2172***	−0.0380***	0.0330***	0.0377***	0.0431***

Panel B: SUE_t

	(A)	(B)	(C)	(D)	(E)
IOF_{t-1}	1.036**			0.875*	
	(0.514)			(0.517)	
$IBuys_{t-1}$		1.075**			0.863*
		(0.521)			(0.524)
$ISales_{t-1}$		−1.024**			−0.880*
		(0.511)			(0.521)
$Return_{t-1}$			0.062**	0.058**	0.058**
			(0.025)	(0.026)	(0.026)
$Sentiment_{t-1}$			0.192	0.194	0.194
			(0.189)	(0.190)	(0.190)
$Volume_{t-1}$			0.053	0.057	0.065
			(0.165)	(0.166)	(0.220)
F-statistic	5.024	2.626	1.723	2.445	1.955

of individual analyst estimates about the average estimate $E[\widehat{ER_i}]$. The narrower the range of estimates the more severe one expects a stock's reaction to an earnings surprise will be.

The SUE scores associated with each announcement come from the I/B/E/S Summary History file. We winsorize the raw SUE scores at the top and bottom 1% to diminish the impact of extreme values. In addition, we require that the earnings release is covered in the Reuters news data. The number of observations is 9,144.

Table 9 reports the determinants of SUE scores on the day of the earnings release. Panel A shows that SUE is correlated positively with returns and news sentiment and negatively with IOF. As in other tables, estimates are from panel regressions with firm fixed effects and volume, defined as the log of total trading volume, returns, sentiment, and IOF or, alternatively, IBuys and ISales. All explanatory variables are measured on the day prior to the earnings announcement. The estimates show that institutional trading predicts the SUE score. Institutions trade in the right direction before an earnings announcement. IBuys predicts a positive earnings surprise whereas ISales predicts a negative surprise.

The earnings announcement results utilize fundamental information that is not directly controlled by the media. This rules out communication between institutions and reporters being purely responsible for our findings.

5.4. Do institutions believe hype?

A large fraction of the public news is generated by the corporations themselves disclosing information that may impact their market values and reporting changes in their financial conditions or operations via press releases. Firms usually sign up for an account with one of the newswire services such as PR Newswire (PRN), Business Wire (BSW), GlobeNewswire, and Marketwire. Newswire services then post press releases on their own websites and also distribute them, typically free of charge, to local and global media outlets, trade magazines, and financial Internet sites. These press releases are delivered to investors directly without processing by information intermediaries. Unlike stories written by Reuters reporters, the press releases do not contain reporters' opinions on the news being disclosed. This feature makes press releases an attractive approach to studying the relationship between institutional trading and hype generated by firms.

Following the adoption of Regulation Fair Disclosure (Reg FD) in October 2000 and of the Sarbanes-Oxley Act in July 2002, corporate press releases became a prevalent method of communicating new developments to investors. The goal of Reg FD is to achieve timeliness and non-exclusivity of corporate information. Communicating information via press releases is a preferred method of accomplishing this.[8] However, under the umbrella of Reg FD, firms may use press releases to hype themselves by releasing positive but potentially controversial or unsubstantiated information to investors in the hope that they will revise their expectations about the firm's value upward and bid up its stock price, at least temporarily.[9] Examples of such news include, but are not limited to, awards for a company's achievements, awards for a product's achievements, securing business from a new or an existing customer, signing new strategic agreements with another firm, reaching milestones or an anniversary, pre-announcement of strong financial results, settlement of litigation against the firm, secured approval to acquire another firm, recruitment or election of top management or board members, promotion, retirement, resignation/departure of top management or board members, sponsorship of an industry event, announcement of an FDA product approval, launch of a new service or introduction of a new product, and improvement or update of a product/service. Indications of opportunistic behavior by firms are returns having little correlation with the sentiment of the press releases and/or sentiment reversing after the original news release. We use press releases as a proxy for "hyping" behavior by corporations to test whether institutions believe the hype.

Reuters provides a comprehensive data set of 76,213 press releases from PRN and 51,073 press releases from BSW matching the time span of our NYSE institutional buy and sell data. Table 10 reports sample statistics for the PRN and BSW releases. For both news sources sentiment has a fat right tail with a fairly large number of extreme positive realizations: both PRN sentiment and BSW sentiment have high average (median) sentiments of 0.290 (0.337) and 0.317 (0.378), respectively, while the average sentiment for Reuters written news is zero. Sentiment distributions also reveal that press releases from both news sources contain a substantial amount of negative news. This is consistent with findings of Neuhierl, Scherbina, and Schlusche (2013), who report significant negative market reaction in response to pre-announcements of disappointing financial results, announcements of FDA rejections, customer losses, product defects, and earnings restatements. In addition, PRN sentiment is weakly positively correlated with returns, 0.0286, and negatively correlated with IOF, -0.0094. BSW sentiment reveals similar correlation coefficients. The sentiment of Reuters written stories is more positively correlated with these variables.

Fig. 5 shows buy-and-hold cumulative stock returns and institutional order flow on press release days for PRN (Panel A) and BSW (Panel B). A ten-day window around the press release day is used in both cases. Buy-and-hold

---

[8] Neuhierl, Scherbina, and Schlusche (2013) study the market reaction to press releases. Unlike our study, they do not discriminate between institutional investors and other investors. They document significant reactions to news about corporate strategy, customers and partners, products and services, management changes, and legal developments.

[9] Here is an example of a typical self-promoting press release by a U. S. corporation courtesy of Neuhierl, Scherbina, and Schlusche (2013). On January 9, 2007 Apple Inc. issued a press release, headlined "Apple Reinvents the Phone with iPhone," which stated: "iPhone … ushers in an era of software power and sophistication never before seen in a mobile device, which completely redefines what users can do on their mobile phones." It contained a pronouncement from CEO Steve Jobs, that "… iPhone is a revolutionary and magical product that is literally five years ahead of any other mobile phone," and described the new product's features. On the day of the announcement Apple's stock price also rose, and in the period from the day before to five days after the announcement the stock earned a cumulative return of 9.31% in excess of the market.

Table 10

Descriptive statistics for hype.

The table reports descriptive statistics for the news data from different news sources. The news sources are PR Newswire (PRN) and Business Wire (BSW). Sentiment is computed as the relevance-weighted average of the difference between positive and negative sentiment scores. Institutional order flow IOF (institutional volume IVol) is defined as the difference between (sum of) institutional purchases IBuys and institutional sales ISales. All trade-related quantities are normalized by the firm's market capitalization lagged by one year and expressed in percent. The sample contains all 1,667 NYSE stocks on hype days from 2003 through 2005. The number of observations is reported in the second column of Panel A. Levels of significance are denoted by * (10%), ** (5%), and *** (1%).

Panel A: Sentiment by news source

	N	Mean	S.D.	5%	50%	95%
PRN sentiment$_t$	76,213	0.290	0.388	−0.476	0.337	0.795
BSW sentiment$_t$	51,073	0.317	0.391	−0.464	0.378	0.796

Panel B: Correlations

	Return$_t$	IOF$_t$	IBuys$_t$	ISales$_t$	Volume$_t$
PRN sentiment$_t$	0.0286***	−0.0094***	−0.0686***	−0.0662***	−0.0640***
BSW sentiment$_t$	0.0291***	−0.0022	−0.0351***	−0.0350***	−0.0310***
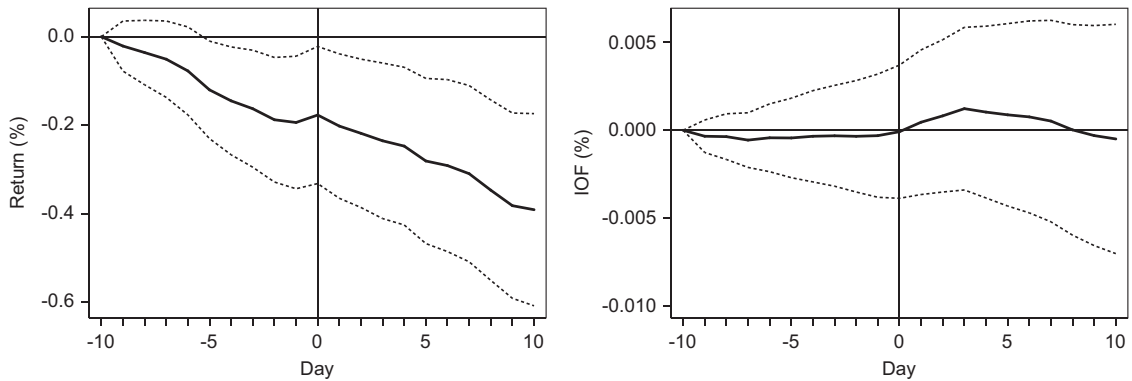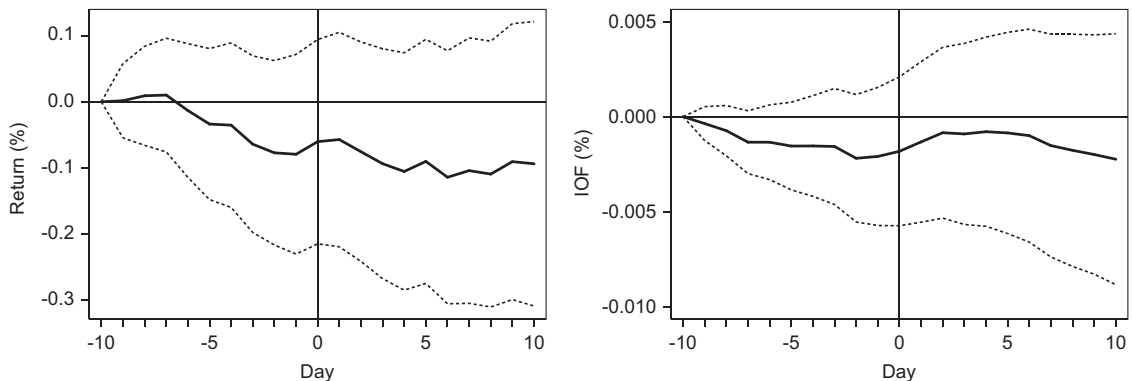
Fig. 5. Hype. The figure reports stock returns and institutional order flow on hype days. In Panel A, hype days are defined as days with PR Newswire (PRN) press releases. In Panel B, hype days are defined as days with Business Wire (BSW) press releases. From left to right, the first plot reports buy-and-hold cumulative stock returns between ten days before and ten days after a news announcement. The second plot reports buy-and-hold cumulative institutional order flow over the same time period. The mean values are calculated as the value-weighted average of the individual news day values. The dotted lines indicate 95% confidence bounds. The sample contains all 1,667 NYSE stocks on hype days from 2003 through 2005. The number of observations is reported in the panel header. Observations are value-weighted. Standard errors are robust to heteroskedasticity and clustering.

cumulative returns provide support for firms strategically using press releases for "hype": returns decline before the announcement day, then show a small jump on the announcement day, and then continue declining. Overall, the pattern is consistent with the hypothesis that firms using press releases are experiencing negative performance and try to temporarily "hype" their stock price by facilitating positive news announcements. Institutional order flow is flat for both news agencies before and right after the announcement day, while rising a little afterward when price declines. Institutional order flow showing no abnormal activity prior to and on the announcement day is consistent with institutions not believing the hype.

Table 11 performs a more thorough investigation of the conjecture that institutions do not believe the hype by using specifications B and E in Table 4. As in other tables, estimates are from panel regressions with firm fixed effects and volume, defined as the log of total trading volume, returns, sentiment, and IBuys and ISales. All explanatory variables are measured on the day prior to the news announcement. The estimates show that institutional trading predicts returns (columns A and B) for both PRN and BSW. However, consistent with the evidence from Fig. 5 institutional trading does not predict

sentiment (columns C and D) for both PRN and BSW. In fact, regression coefficients for IBuys and ISales for the PRN category are both negative, albeit not statistically significant. For BSW these coefficients are not statistically significant. These results provide further support for the conjecture that institutions do not believe the hype.

We next make use of the Reuters news data to provide additional evidence on whether institutions believe the hype. In the Reuters news data we select news events with subsequent sentiment reversal over five and ten days. Out of these news events we look at days with the 10% largest sentiment reversals. These news events can be interpreted either as hype or as transitory or erroneous news.

Fig. 6 shows the buy-and-hold cumulative returns and institutional order flows for news days with five-day sentiment reversal (Panel A) and ten-day sentiment reversal (Panel B). In both cases the market responds in the direction of the news: investors earn positive (negative) abnormal returns on the announcement day which declines (increases) on subsequent days. Therefore, sentiment reversals are associated with transitory stock returns. The institutions on average sell before positive news releases and start buying after prices come down after the news releases in the case of five-day sentiment reversals. Before

Table 11
Returns, news sentiment, and institutional trading on hype days.

The table documents the predictability of news announcement day returns (columns A–B), news sentiment (columns C–D), and institutional order flow (columns E–F) on hype days. In Panel A, hype days are defined as days with PR Newswire (PRN) press releases. In Panel B, hype days are defined as days with Business Wire (BSW) press releases. Estimates are from panel regressions with firm fixed effects. IOF denotes institutional order flow, IBuys (ISales) are institutional purchases (sales), Sentiment is the news sentiment, and Volume is the log of total trading volume. All explanatory variables are measured on the day prior to the news announcement. The sample contains all 1,667 NYSE stocks on hype days from 2003 through 2005. The number of observations is reported in the panel header. Observations are value-weighted. Standard errors are robust to heteroskedasticity and clustering and reported in parentheses. Levels of significance are denoted by * (10%), ** (5%), and *** (1%).

	$Return_t$		$Sentiment_t$		IOF_t	
	(A)	(B)	(C)	(D)	(E)	(F)
Panel A: PRN press releases (N=76,213)						
IOF_{t-1}	0.461***		0.008		0.234***	
	(0.164)		(0.028)		(0.013)	
$IBuys_{t-1}$		0.441***		−0.002		0.237***
		(0.165)		(0.029)		(0.013)
$ISales_{t-1}$		−0.486***		−0.021		−0.231***
		(0.165)		(0.027)		(0.013)
$Return_{t-1}$	−0.028*	−0.028*	0.003	0.003	0.004***	0.004***
	(0.016)	(0.016)	(0.003)	(0.003)	(0.000)	(0.000)
$Sentiment_{t-1}$	0.044	0.044	0.039**	0.039**	−0.002**	−0.002***
	(0.043)	(0.043)	(0.015)	(0.015)	(0.001)	(0.001)
$Volume_{t-1}$	−0.018	0.000	−0.027***	−0.018*	0.002**	−0.000
	(0.036)	(0.046)	(0.008)	(0.010)	(0.001)	(0.002)
F-statistic	5.540	4.868	7.219	18.113	234.389	193.023
Panel B: BSW press releases (N=51,073)						
IOF_{t-1}	0.996***		0.034		0.246***	
	(0.335)		(0.042)		(0.011)	
$IBuys_{t-1}$		0.950***		0.032		0.244***
		(0.340)		(0.042)		(0.011)
$ISales_{t-1}$		−1.054***		−0.037		−0.248***
		(0.337)		(0.043)		(0.010)
$Return_{t-1}$	−0.027	−0.028	−0.001	−0.001	0.004***	0.004***
	(0.022)	(0.021)	(0.002)	(0.002)	(0.000)	(0.000)
$Sentiment_{t-1}$	0.044	0.044	0.037**	0.037**	−0.001	−0.001
	(0.039)	(0.039)	(0.015)	(0.015)	(0.001)	(0.001)
$Volume_{t-1}$	−0.004	0.033	−0.005	−0.003	0.001	0.002
	(0.043)	(0.055)	(0.010)	(0.013)	(0.001)	(0.002)
F-statistic	7.800	6.833	1.439	1.323	205.310	192.123

A

5-Day sentiment reversal (*N*=12,563)

B

10-Day sentiment reversal (*N*= 12,447)
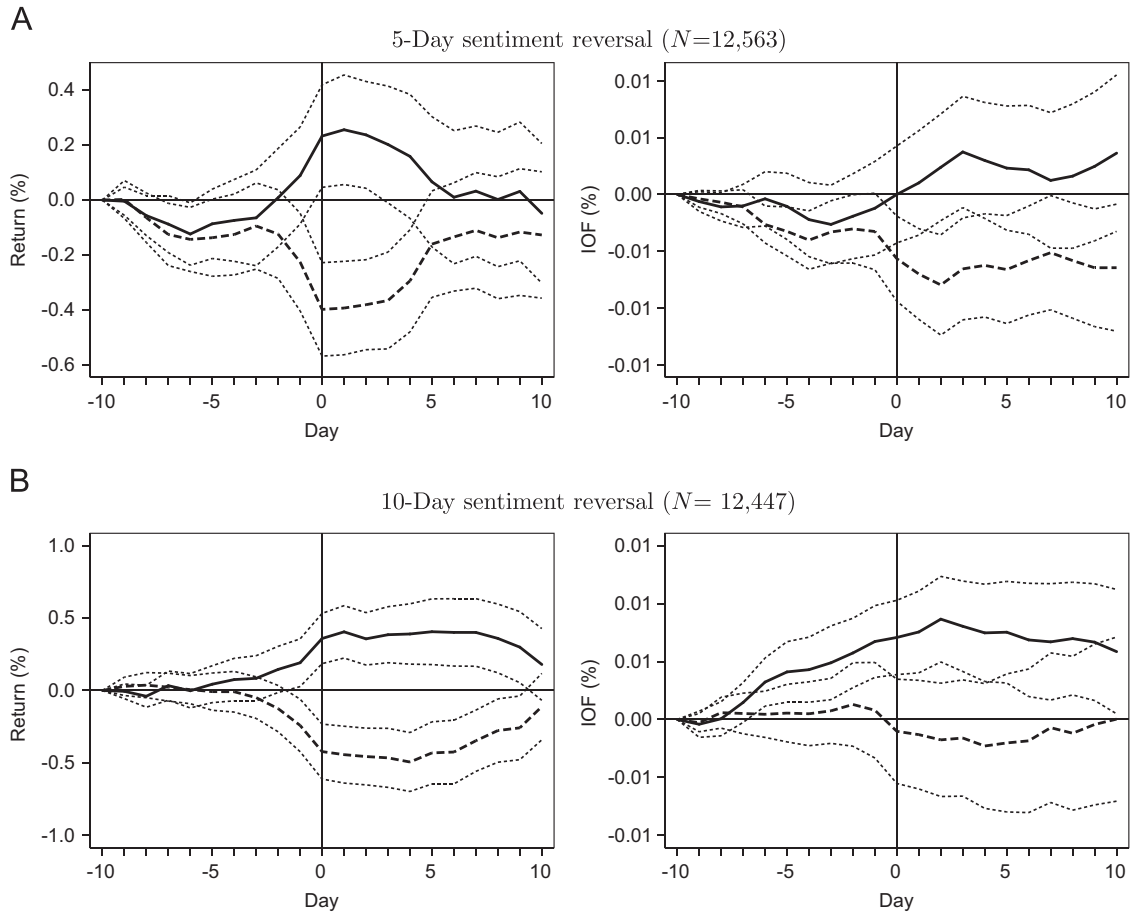
Fig. 6. Sentiment reversals. The figure reports sentiment, return, and institutional order flow around sentiment reversal days. In Panel A, sentiment reversal days are defined as days with news in either direction followed by news in the opposite direction within five days. In Panel B, sentiment reversal days are defined as days with news in either direction followed by news in the opposite direction within ten days. We look at the top 10% of days with largest sentiment reversals. From left to right, the first plot reports buy-and-hold cumulative stock returns between ten days before and ten days after a news announcement. The second plot reports buy-and-hold cumulative institutional order flow over the same time period. The solid line represents good news days, defined as the top sentiment quintile on news announcement days. The dashed line represents bad news days, defined as the bottom sentiment quintile on news announcement days. The mean values are calculated as the value-weighted average of the individual news day values. The dotted lines indicate 95% confidence bounds. The sample contains all 1,667 NYSE stocks on sentiment reversal days from 2003 through 2005. The number of observations is reported in the panel header. Observations are value-weighted. Standard errors are robust to heteroskedasticity and clustering.

and after negative news releases, the institutions on average sell without any sign of a significant order flow on the announcement day in the case of five-day sentiment reversals. Institutions tend to buy before positive news releases and keep buying after prices come down after the news releases in the case of ten-day sentiment reversals. However, this result is not statistically significant. The institutions on average do not show significant market activity either before or after negative news reversals.

Table 12 establishes a quantitative relation between sentiment, returns, and institutional trading using the same regression specification as in Table 11. As in other tables, estimates are from panel regressions with firm fixed effects and volume, defined as the log of total trading volume, returns, sentiment, and IBuys and ISales. All explanatory variables are measured on the day prior to the news announcement. The estimates show that institutional trading does not have predictive power, neither for returns nor sentiment, on news days with five-day reversals (Panel A). The regression coefficients for IBuys and ISales have wrong

signs in the sentiment regressions. On news days with ten-day reversals (Panel B) institutional trading does predict returns but not sentiment. While the regression coefficients for IBuys and ISales have the correct signs in the sentiment regression, they are not statistically significant. For BSW these coefficients are not statistically significant.

As in the Martha Stewart example, the results from press releases and news reversals support the conjecture that institutions do not believe hype in the media, but rather act as long-term informed investors.

6. Conclusion

This paper combines daily non-public data on buy and sell volume by institutions with news announcements from Reuters. Natural language processing categorizes the sentiment associated with each news story. We find that institutional trading predicts news announcements, the sentiment of the news, returns on announcement day, and earnings

Table 12

Returns, news sentiment, and institutional trading on sentiment reversal days.

The table documents the predictability of news announcement day returns (columns A–B), news sentiment (columns C–D), and institutional order flow (columns E–F) on sentiment reversal days. In Panel A, sentiment reversal days are defined as days with news in either direction followed by news in the opposite direction within five days. In Panel B, sentiment reversal days are defined as days with news in either direction followed by news in the opposite direction within ten days. We look at the top 10% of days with largest sentiment reversals. Estimates are from panel regressions with firm fixed effects. IOF denotes institutional order flow, IBuys (ISales) are institutional purchases (sales), Sentiment is the news sentiment, and Volume is the log of total trading volume. All explanatory variables are measured on the day prior to the news announcement. The sample contains all 1,667 NYSE stocks on sentiment reversal days from 2003 through 2005. The number of observations is reported in the panel header. Observations are value-weighted. Standard errors are robust to heteroskedasticity and clustering and reported in parentheses. Levels of significance are denoted by * (10%), ** (5%), and *** (1%).

	$Return_t$		$Sentiment_t$		IOF_t	
	(A)	(B)	(C)	(D)	(E)	(F)
Panel A: 5-Day sentiment reversal (N=12,563)						
IOF_{t-1}	0.479		−0.030		0.225***	
	(0.330)		(0.076)		(0.023)	
$IBuys_{t-1}$		0.495		−0.033		0.226***
		(0.332)		(0.076)		(0.021)
$ISales_{t-1}$		−0.450		0.025		−0.224***
		(0.328)		(0.077)		(0.026)
$Return_{t-1}$	−0.023	−0.023	0.013*	0.013*	0.004***	0.004***
	(0.025)	(0.025)	(0.007)	(0.007)	(0.001)	(0.001)
$Sentiment_{t-1}$	0.033	0.033	0.208***	0.208***	−0.001	−0.001
	(0.083)	(0.083)	(0.023)	(0.023)	(0.001)	(0.001)
$Volume_{t-1}$	−0.092	−0.113	−0.034	−0.031	0.001	0.001
	(0.065)	(0.078)	(0.022)	(0.029)	(0.002)	(0.003)
Constant	0.009	−0.008	−0.029***	−0.026**	0.002***	0.001
	(0.036)	(0.046)	(0.006)	(0.012)	(0.001)	(0.002)
F-statistic	1.814	1.532	14.409	13.212	42.781	34.348
Panel B: 10-Day sentiment reversal (N=12,447)						
IOF_{t-1}	0.839***		0.063		0.236***	
	(0.265)		(0.071)		(0.016)	
$IBuys_{t-1}$		0.850***		0.068		0.237***
		(0.264)		(0.071)		(0.016)
$ISales_{t-1}$		−0.802***		−0.048		−0.234***
		(0.272)		(0.071)		(0.018)
$Return_{t-1}$	−0.014	−0.014	0.011*	0.011*	0.004***	0.004***
	(0.025)	(0.025)	(0.006)	(0.006)	(0.001)	(0.001)
$Sentiment_{t-1}$	0.266***	0.265***	0.254***	0.254***	−0.001	−0.001
	(0.084)	(0.084)	(0.027)	(0.027)	(0.002)	(0.002)
$Volume_{t-1}$	−0.039	−0.063	−0.068***	−0.078***	0.004**	0.003
	(0.065)	(0.077)	(0.019)	(0.024)	(0.002)	(0.002)
Constant	0.041	0.022	−0.026***	−0.034***	0.003***	0.001
	(0.034)	(0.042)	(0.005)	(0.009)	(0.001)	(0.002)
F-statistic	6.151	5.157	22.681	18.963	71.970	58.053

announcement surprises. These findings suggest that institutions are producing value-relevant information for stocks and support the findings based on institutional holdings and trading that institutions improve price efficiency (Badrinath, Kale, and Noe, 1995; Sias and Starks, 1997; Campbell, Ramadorai, and Schwartz, 2009; Boehmer and Kelley, 2009).[10] Our results also provide direct evidence on Tetlock's (2010) finding that news reduces informational asymmetry.

Regarding specific news types we find that while institutions have price impact on macroeconomic news days they are not informed about this type of news, with

the exception of news on economic indicators. On the other hand, institutions are informed about a range of unexpected value-destroying events as well as earnings announcement surprises. Finally, we show that institutions do not believe the media hype as defined by press releases and news stories that are reversed.

The source of institutions' informational advantage is difficult to know. Their ability to predict news regarding earnings could arise from the ability to better process public information or from sources that should be proscribed by Reg FD. However, institutions' ability to forecast unexpected value-destroying events provides evidence in favor of the ability to better process public information. Further research on the sources of institutions' informational advantages will continue to be important.

An important issue raised by our results is how to evaluate what constitutes true news. If some investors trade on the news in advance of its publication, then evidence showing a weak link between news and stock

---

[10] Prior literature using other measures of institutional trading finds limited evidence supporting institutions being informed (Griffin, Shu, and Topaloglu, 2012; Jegadeesh and Tang, 2010; Busse, Green, and Jegadeesh, 2012). One explanation is that broader, more comprehensive measures of institutional data are needed (as also in Campbell, Ramadorai, and Schwartz, 2009).
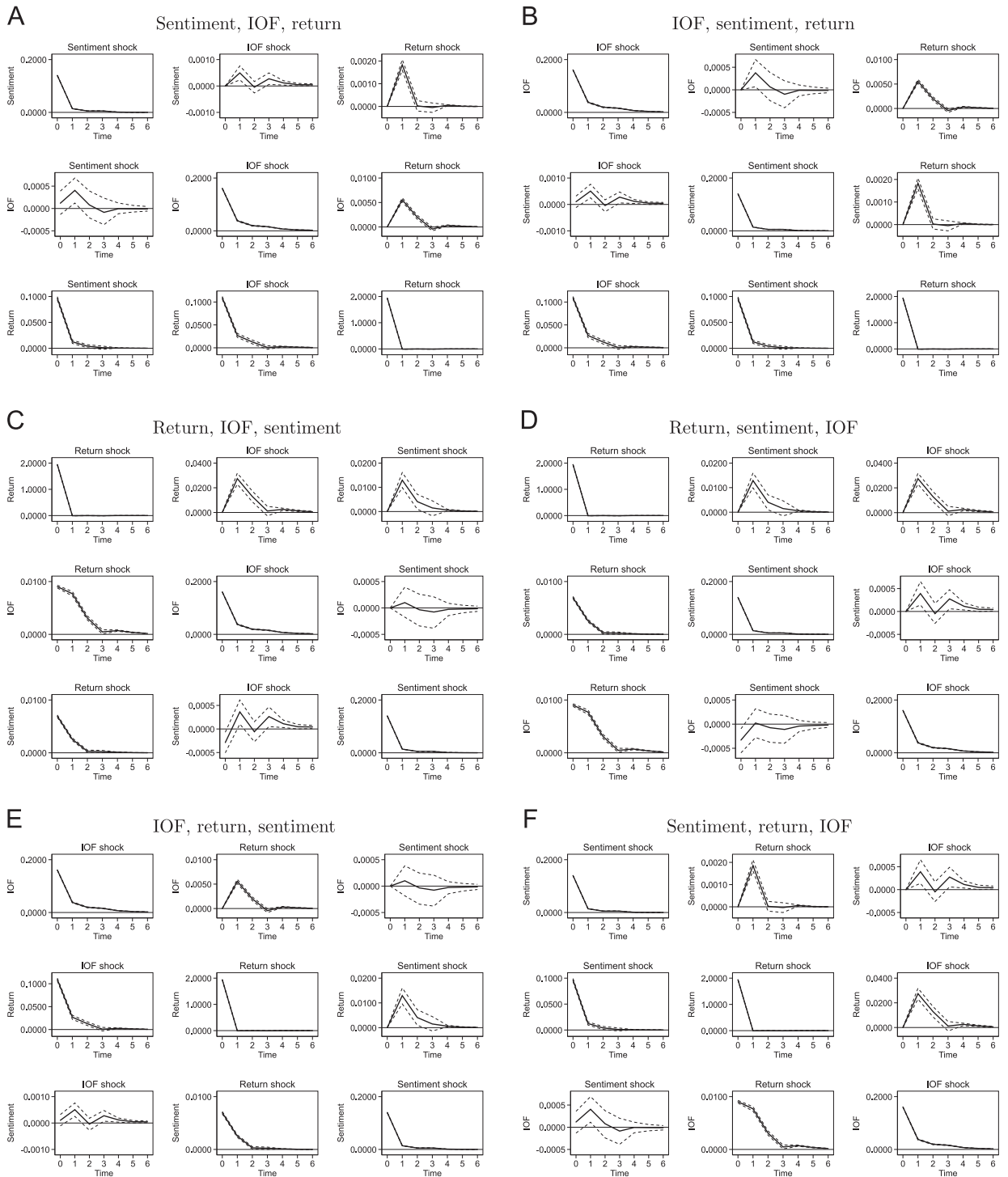
Fig. A1. Impulse responses. The figure reports the impulse response functions corresponding to the panel VAR with lag length $L=3$ reported in Table 6. The estimates in Table 6 are obtained using GMM estimation as described in Section 4.2. Impulse responses correspond to a one standard deviation shock. Error bands at 5% level for the impulse responses (dashed lines) are generated using Monte-Carlo simulations with 1,000 draws. Across panels, the dependent variables are ordered in varying sequences: (A) Sentiment, IOF, return; (B) IOF, sentiment, return; (C) Return, IOF, sentiment; (D) Return, sentiment, IOF; (E) IOF, return, sentiment; and (F) Sentiment, return, IOF.

prices, e.g., Roll's (1988) presidential address and Boudoukh, Feldman, Kogan, and Richardson (2013), must be interpreted with caution. A common approach

posits that individual stocks' R^2 from factor models should differ across days with news and without news. However, this requires the econometrician to observe

news arrival. If, as our results suggest, some of the information content in the news is incorporated prior to the release date, then the ability to measure days with news and days without news is questionable. This highlights the importance of further study of specific institutions' trading and information flows to help better understand the informational efficiency of stock prices and the statistical properties of returns.

Appendix A

See Fig. A1, Tables A1–A10.

Table A1
Returns, news sentiment, and institutional trading at industry and market level.

The table documents the predictability of news announcement day returns (columns A–B), news sentiment (columns C–D), and institutional order flow (columns E–F) at industry and market level. In Panel A, all variables are value-weighted aggregates at the industry level. We use the Fama-French definition for 12 industries, with 1=NoDur, 2=Durbl, 3=Manuf, 4=Enrgy, 5=Chems, 6=BusEq, 7=Telcm, 8=Utils, 9=Shops, 10=Hlth, 11=Money, 12=Other. In Panel B, all variables are value-weighted aggregates at the market level. Estimates are from pooled OLS regressions. IOF denotes institutional order flow, IBuys (ISales) are institutional purchases (sales), Sentiment is the news sentiment, and Volume is the log of total trading volume. All explanatory variables are measured on the day prior to the news announcement. The sample contains all 1,667 NYSE stocks on news days from 2003 through 2005. The number of observations is reported in the panel header. Observations are value-weighted. Standard errors are robust to heteroskedasticity and clustering and reported in parentheses. Levels of significance are denoted by * (10%), ** (5%), and *** (1%).

	$Return_t$		$Sentiment_t$		IOF_t	
	(A)	(B)	(C)	(D)	(E)	(F)
Panel A: Industry (N=8,884)						
IOF_{t-1}	2.736*		0.531**		0.259***	
	(1.406)		(0.261)		(0.024)	
$IBuys_{t-1}$		2.588*		0.509*		0.259***
		(1.467)		(0.279)		(0.025)
$ISales_{t-1}$		−2.864**		−0.549**		−0.258***
		(1.360)		(0.255)		(0.023)
$Return_{t-1}$	−0.013	−0.015	0.008***	0.008***	0.004***	0.004***
	(0.024)	(0.024)	(0.003)	(0.003)	(0.000)	(0.000)
$Sentiment_{t-1}$	−0.023	−0.025	0.366***	0.366***	0.002**	0.002**
	(0.092)	(0.093)	(0.058)	(0.058)	(0.001)	(0.001)
$Volume_{t-1}$	−0.032	0.047	0.004	0.015	0.001	0.001
	(0.047)	(0.115)	(0.024)	(0.027)	(0.001)	(0.002)
F-statistic	2.694	2.197	52.574	42.026	87.152	69.851
Panel B: Market (N=742)						
IOF_{t-1}	−1.850		0.576**		0.276***	
	(3.359)		(0.268)		(0.051)	
$IBuys_{t-1}$		−1.996		0.653**		0.288***
		(3.385)		(0.272)		(0.050)
$ISales_{t-1}$		1.312		−0.296		−0.234***
		(3.494)		(0.278)		(0.054)
$Return_{t-1}$	−0.100*	−0.100*	0.010**	0.010**	0.005***	0.005***
	(0.058)	(0.058)	(0.005)	(0.005)	(0.001)	(0.001)
$Sentiment_{t-1}$	0.332	0.453	0.665***	0.602***	0.005	−0.004
	(0.918)	(0.902)	(0.077)	(0.075)	(0.011)	(0.012)
$Volume_{t-1}$	−0.195	−0.034	0.031	−0.053	0.003	−0.009**
	(0.148)	(0.358)	(0.020)	(0.048)	(0.002)	(0.005)
F-statistic	1.430	1.230	31.443	27.039	14.999	14.413

Table A2
Returns, news sentiment, and institutional trading on announcement days sorted by size.

The table documents the predictability of news announcement returns (Panel A), news sentiment (Panel B), and institutional order flow (Panel C) on news days. Estimates are from panel regressions with firm fixed effects. IOF denotes institutional order flow, IBuys (ISales) are institutional purchases (sales), Sentiment is the news sentiment, and Volume is the log of total trading volume. All explanatory variables are measured on the day prior to the news announcement. The sample contains all 1,667 NYSE stocks on news days from 2003 through 2005. We sort all firms into size terciles. The number of observations is 42,049 in each portfolio. Observations are value-weighted. Standard errors are robust to heteroskedasticity and clustering and reported in parentheses. Levels of significance are denoted by * (10%), ** (5%), and *** (1%).

	Firm size					
	Small		Medium		Large	
	(A)	(B)	(C)	(D)	(E)	(F)
Panel A: Return$_t$						
IOF$_{t-1}$	0.291***		0.355**		1.434***	
	(0.077)		(0.145)		(0.434)	
IBuys$_{t-1}$		0.278***		0.351**		1.326***
		(0.078)		(0.145)		(0.431)
ISales$_{t-1}$		−0.302***		−0.364**		−1.605***
		(0.078)		(0.147)		(0.451)
Return$_{t-1}$	0.005	0.005	−0.017	−0.017	−0.016	−0.018
	(0.013)	(0.013)	(0.013)	(0.013)	(0.016)	(0.016)
Sentiment$_{t-1}$	0.001	0.003	0.064*	0.064*	0.030	0.029
	(0.066)	(0.066)	(0.036)	(0.036)	(0.025)	(0.025)
Volume$_{t-1}$	−0.065**	−0.039	−0.024	−0.014	−0.023	0.053
	(0.031)	(0.038)	(0.027)	(0.036)	(0.033)	(0.055)
F-statistic	6.340	5.333	4.481	3.585	11.133	10.133
Panel B: Sentiment$_t$						
IOF$_{t-1}$	0.024**		−0.011		0.147***	
	(0.011)		(0.022)		(0.055)	
IBuys$_{t-1}$		0.020*		−0.012		0.113**
		(0.011)		(0.022)		(0.055)
ISales$_{t-1}$		−0.028**		0.006		−0.200***
		(0.011)		(0.022)		(0.060)
Return$_{t-1}$	0.011***	0.011***	0.009***	0.009***	0.008***	0.008***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Sentiment$_{t-1}$	0.296***	0.297***	0.259***	0.259***	0.154***	0.154***
	(0.022)	(0.022)	(0.015)	(0.015)	(0.011)	(0.012)
Volume$_{t-1}$	−0.007	0.002	−0.010*	−0.005	−0.015	0.009
	(0.004)	(0.005)	(0.006)	(0.007)	(0.010)	(0.012)
F-statistic	191.135	154.712	250.011	200.369	127.429	116.738
Panel C: IOF$_t$						
IOF$_{t-1}$	0.219***		0.277***		0.235***	
	(0.016)		(0.013)		(0.016)	
IBuys$_{t-1}$		0.224***		0.278***		0.239***
		(0.016)		(0.013)		(0.017)
ISales$_{t-1}$		−0.215***		−0.273***		−0.229***
		(0.016)		(0.015)		(0.016)
Return$_{t-1}$	0.006***	0.006***	0.005***	0.005***	0.003***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
Sentiment$_{t-1}$	0.006	0.005	−0.001	−0.001	−0.001**	−0.001**
	(0.005)	(0.005)	(0.002)	(0.002)	(0.000)	(0.000)
Volume$_{t-1}$	0.000	−0.010***	0.002	−0.002	0.002**	−0.001
	(0.003)	(0.003)	(0.002)	(0.003)	(0.001)	(0.001)
F-statistic	89.583	79.640	192.454	164.567	304.217	250.351

Table A3

Returns, news sentiment, and institutional trading on announcement days sorted by number of analysts.

The table documents the predictability of news announcement returns (Panel A), news sentiment (Panel B), and institutional order flow (Panel C) on news days. Estimates are from panel regressions with firm fixed effects. IOF denotes institutional order flow, IBuys (ISales) are institutional purchases (sales), Sentiment is the news sentiment, and Volume is the log of total trading volume. All explanatory variables are measured on the day prior to the news announcement. The sample contains all 1,667 NYSE stocks on news days from 2003 through 2005. We sort all firms into terciles. The number of observations is 42,049 in each portfolio. Observations are value-weighted. Standard errors are robust to heteroskedasticity and clustering and reported in parentheses. Levels of significance are denoted by * (10%), ** (5%), and *** (1%).

| | No. analysts | | | | | |
| | Small | | Medium | | Large | |
	(A)	(B)	(C)	(D)	(E)	(F)
Panel A: $Return_t$						
IOF_{t-1}	0.324***		0.854***		0.896***	
	(0.097)		(0.308)		(0.308)	
$IBuys_{t-1}$		0.318***		0.823***		0.857***
		(0.101)		(0.309)		(0.302)
$ISales_{t-1}$		−0.328***		−0.912***		−0.967***
		(0.095)		(0.308)		(0.323)
$Return_{t-1}$	−0.006	−0.006	−0.021	−0.021	−0.015	−0.016
	(0.026)	(0.026)	(0.016)	(0.016)	(0.016)	(0.016)
$Sentiment_{t-1}$	0.083	0.083	0.028	0.028	0.022	0.022
	(0.073)	(0.073)	(0.051)	(0.051)	(0.026)	(0.026)
$Volume_{t-1}$	−0.010	−0.004	−0.071*	−0.036	−0.002	0.034
	(0.026)	(0.031)	(0.040)	(0.049)	(0.032)	(0.044)
F-statistic	2.514	2.086	9.786	9.591	5.573	4.940
Panel B: $Sentiment_t$						
IOF_{t-1}	0.024		0.043		0.076	
	(0.025)		(0.033)		(0.047)	
$IBuys_{t-1}$		0.023		0.033		0.067
		(0.025)		(0.032)		(0.046)
$ISales_{t-1}$		−0.025		−0.061*		−0.092*
		(0.024)		(0.036)		(0.049)
$Return_{t-1}$	0.012***	0.012***	0.008***	0.008***	0.009***	0.009***
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
$Sentiment_{t-1}$	0.215***	0.215***	0.178***	0.178***	0.150***	0.150***
	(0.031)	(0.031)	(0.013)	(0.013)	(0.015)	(0.015)
$Volume_{t-1}$	−0.015*	−0.014	0.003	0.014	−0.022**	−0.014
	(0.009)	(0.010)	(0.011)	(0.013)	(0.010)	(0.013)
F-statistic	37.535	31.173	75.957	64.384	82.235	74.208
Panel C: IOF_t						
IOF_{t-1}	0.224***		0.245***		0.254***	
	(0.020)		(0.012)		(0.016)	
$IBuys_{t-1}$		0.232***		0.247***		0.256***
		(0.020)		(0.012)		(0.017)
$ISales_{t-1}$		−0.219***		−0.242***		−0.250***
		(0.021)		(0.012)		(0.016)
$Return_{t-1}$	0.004***	0.004***	0.004***	0.004***	0.003***	0.003***
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
$Sentiment_{t-1}$	−0.001	−0.001	−0.001	−0.001	−0.001**	−0.001**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$Volume_{t-1}$	0.004**	−0.004	0.004***	0.001	0.001	−0.001
	(0.002)	(0.003)	(0.001)	(0.002)	(0.001)	(0.002)
F-statistic	78.077	70.434	272.953	220.464	213.187	172.930

Table A4
Returns, news sentiment, and institutional trading on announcement days sorted by number of firms affected by news.

    The table documents the predictability of news announcement returns (Panel A), news sentiment (Panel B), and institutional order flow (Panel C) on news days. Estimates are from panel regressions with firm fixed effects. IOF denotes institutional order flow, IBuys (ISales) are institutional purchases (sales), Sentiment is the news sentiment, and Volume is the log of total trading volume. All explanatory variables are measured on the day prior to the news announcement. The sample contains all 1,667 NYSE stocks on news days from 2003 through 2005. We sort all news into terciles. The number of observations is 42,049 in each portfolio. Observations are value-weighted. Standard errors are robust to heteroskedasticity and clustering and reported in parentheses. Levels of significance are denoted by * (10%), ** (5%), and *** (1%).

| | No. firms in news story | | | | | |
| | Small | | Medium | | Large | |
	(A)	(B)	(C)	(D)	(E)	(F)
Panel A: Return$_t$						
IOF$_{t-1}$	0.405***		0.740***		0.917***	
	(0.142)		(0.196)		(0.327)	
IBuys$_{t-1}$		0.394***		0.718***		0.870***
		(0.142)		(0.197)		(0.328)
ISales$_{t-1}$		−0.418***		−0.793***		−0.977***
		(0.145)		(0.206)		(0.328)
Return$_{t-1}$	−0.021	−0.022	−0.024	−0.025	−0.011	−0.011
	(0.014)	(0.014)	(0.016)	(0.016)	(0.017)	(0.017)
Sentiment$_{t-1}$	0.027	0.028	0.040	0.041	0.030	0.030
	(0.047)	(0.047)	(0.047)	(0.047)	(0.035)	(0.035)
Volume$_{t-1}$	0.030	0.040	−0.006	0.024	−0.053	−0.015
	(0.035)	(0.043)	(0.042)	(0.050)	(0.033)	(0.042)
F-statistic	5.398	4.313	8.553	7.420	6.404	6.447
Panel B: Sentiment$_t$						
IOF$_{t-1}$	0.023		0.068*		0.070**	
	(0.029)		(0.035)		(0.034)	
IBuys$_{t-1}$		0.022		0.064*		0.059*
		(0.029)		(0.035)		(0.034)
ISales$_{t-1}$		−0.025		−0.078**		−0.085**
		(0.029)		(0.036)		(0.035)
Return$_{t-1}$	0.004	0.004	0.008***	0.008***	0.010***	0.010***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.001)	(0.001)
Sentiment$_{t-1}$	0.183***	0.183***	0.160***	0.160***	0.152***	0.152***
	(0.014)	(0.014)	(0.015)	(0.015)	(0.013)	(0.013)
Volume$_{t-1}$	−0.015**	−0.014	0.003	0.009	−0.015*	−0.006
	(0.008)	(0.010)	(0.009)	(0.011)	(0.009)	(0.010)
F-statistic	34.496	28.617	36.667	30.031	99.929	92.302
Panel C: IOF$_t$						
IBuys$_{t-1}$		0.251***		0.244***		0.250***
		(0.020)		(0.014)		(0.015)
ISales$_{t-1}$		−0.246***		−0.239***		−0.242***
		(0.020)		(0.015)		(0.015)
IOF$_{t-1}$	0.249***		0.242***		0.246***	
	(0.020)		(0.015)		(0.015)	
Return$_{t-1}$	0.004***	0.004***	0.004***	0.004***	0.003***	0.003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Sentiment$_{t-1}$	−0.000	−0.000	0.001	0.001	−0.002***	−0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Volume$_{t-1}$	0.000	−0.002	−0.000	−0.002	0.003***	0.000
	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)
F-statistic	212.574	175.816	184.025	153.178	190.596	157.004

Table A5

Returns, news sentiment, and institutional trading on announcement days sorted by the GIM governance index.

The table documents the predictability of news announcement returns (Panel A), news sentiment (Panel B), and institutional order flow (Panel C) on news days. Estimates are from panel regressions with firm fixed effects. IOF denotes institutional order flow, IBuys (ISales) are institutional purchases (sales), Sentiment is the news sentiment, and Volume is the log of total trading volume. All explanatory variables are measured on the day prior to the news announcement. The sample contains all 1,667 NYSE stocks on news days from 2003 through 2005. We sort all firms into terciles. The number of observations is 42,049 in each portfolio. Observations are value-weighted. Standard errors are robust to heteroskedasticity and clustering and reported in parentheses. Levels of significance are denoted by * (10%), ** (5%), and *** (1%).

| | GIM governance index | | | | | |
| | Low | | Medium | | High | |
	(A)	(B)	(C)	(D)	(E)	(F)
Panel A: $Return_t$						
IOF_{t-1}	1.036***		0.506**		0.411*	
	(0.324)		(0.244)		(0.242)	
$IBuys_{t-1}$		1.040***		0.447*		0.392
		(0.327)		(0.238)		(0.253)
$ISales_{t-1}$		−1.033***		−0.583**		−0.438*
		(0.322)		(0.255)		(0.239)
$Return_{t-1}$	−0.031*	−0.031*	0.007	0.006	−0.019	−0.019
	(0.016)	(0.016)	(0.018)	(0.018)	(0.017)	(0.017)
$Sentiment_{t-1}$	−0.007	−0.007	0.016	0.016	0.019	0.020
	(0.038)	(0.038)	(0.047)	(0.047)	(0.071)	(0.071)
$Volume_{t-1}$	0.004	0.001	−0.057	0.003	−0.042	−0.015
	(0.039)	(0.045)	(0.042)	(0.056)	(0.037)	(0.049)
F-statistic	8.953	7.166	2.223	3.735	1.913	1.923
Panel B: $Sentiment_t$						
IOF_{t-1}	0.033		0.050		0.095**	
	(0.040)		(0.045)		(0.039)	
$IBuys_{t-1}$		0.026		0.042		0.095**
		(0.040)		(0.045)		(0.040)
$ISales_{t-1}$		−0.039		−0.061		−0.096**
		(0.040)		(0.046)		(0.040)
$Return_{t-1}$	0.011***	0.011***	0.009***	0.009***	0.010***	0.010***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
$Sentiment_{t-1}$	0.132***	0.132***	0.172***	0.172***	0.163***	0.163***
	(0.016)	(0.016)	(0.014)	(0.014)	(0.024)	(0.024)
$Volume_{t-1}$	0.003	0.007	−0.019*	−0.010	−0.014	−0.013
	(0.011)	(0.013)	(0.011)	(0.014)	(0.012)	(0.015)
F-statistic	44.476	36.165	54.127	49.035	24.386	19.694
Panel C: IOF_t						
IOF_{t-1}	0.244***		0.269***		0.286***	
	(0.015)		(0.014)		(0.021)	
$IBuys_{t-1}$		0.247***		0.268***		0.290***
		(0.016)		(0.014)		(0.019)
$ISales_{t-1}$		−0.242***		−0.269***		−0.282***
		(0.015)		(0.015)		(0.022)
$Return_{t-1}$	0.004***	0.004***	0.004***	0.004***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
$Sentiment_{t-1}$	−0.001	−0.001	−0.002**	−0.002**	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$Volume_{t-1}$	0.002*	0.000	0.001	0.001	0.004**	−0.001
	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.004)
F-statistic	183.433	153.362	179.976	144.558	97.898	81.914

Table A6

Returns, news sentiment, and institutional trading on announcement days sorted by industry.

The table documents the predictability of news announcement returns (Panel A), news sentiment (Panel B), and institutional order flow (Panel C) on news days. Estimates are from panel regressions with firm fixed effects. IOF denotes institutional order flow, IBuys (ISales) are institutional purchases (sales), Sentiment is the news sentiment, and Volume is the log of total trading volume. All explanatory variables are measured on the day prior to the news announcement. The sample contains all 1,667 NYSE stocks on news days from 2003 through 2005. We sort all firms into the 12 Fama-French industries. The number of observations in each portfolio is reported in the second row of the table. Observations are value-weighted. Standard errors are robust to heteroskedasticity and clustering and reported in parentheses. Levels of significance are denoted by * (10%), ** (5%), and *** (1%).

	1 NoDur	2 Durbl	3 Manuf	4 Enrgy	5 Chems	6 BusEq	7 Telcm	8 Utils	9 Shops	10 Hlth	11 Money	12 Other
N	9,523	4,026	12,998	8,219	4,358	11,172	3,953	12,312	13,981	10,121	21,717	13,768
Panel A: Return$_t$												
IBuys$_{t-1}$	0.562*	2.785	0.539	1.214***	0.780	0.466	1.310**	0.605*	−0.071	0.802	0.771**	0.431*
	(0.305)	(1.779)	(0.334)	(0.439)	(0.481)	(0.354)	(0.604)	(0.325)	(0.339)	(0.513)	(0.301)	(0.221)
ISales$_{t-1}$	−0.559*	−2.890*	−0.577*	−1.273***	−0.900*	−0.473	−1.322**	−0.670*	−0.016	−0.914	−0.884***	−0.485**
	(0.296)	(1.745)	(0.325)	(0.451)	(0.484)	(0.386)	(0.540)	(0.360)	(0.326)	(0.590)	(0.326)	(0.222)
Return$_{t-1}$	−0.003	−0.014	−0.038	−0.027	−0.013	−0.002	−0.026*	−0.013	−0.044**	0.017	−0.019	−0.037
	(0.039)	(0.013)	(0.026)	(0.034)	(0.027)	(0.016)	(0.015)	(0.029)	(0.019)	(0.024)	(0.018)	(0.023)
Sentiment$_{t-1}$	0.140*	−0.178***	0.036	0.055	−0.032	0.019	0.199***	0.015	0.012	−0.033	0.096**	−0.015
	(0.084)	(0.037)	(0.051)	(0.076)	(0.060)	(0.060)	(0.054)	(0.054)	(0.060)	(0.043)	(0.040)	(0.074)
Volume$_{t-1}$	−0.031	−0.058	−0.015	−0.007	0.026	−0.019	−0.149***	−0.004	0.072	0.034	0.007	0.039
	(0.044)	(0.040)	(0.058)	(0.081)	(0.062)	(0.064)	(0.042)	(0.058)	(0.067)	(0.079)	(0.052)	(0.051)
F-statistic	1.292	3.666	3.024	3.577	0.610	0.613	2.383	1.729	1.798	1.019	5.046	2.845
Panel B: Sentiment$_t$												
IBuys$_{t-1}$	−0.005	0.019	0.077	−0.074	0.009	0.006	0.014	−0.053	0.052	0.164	0.052	0.080**
	(0.105)	(0.078)	(0.049)	(0.057)	(0.219)	(0.043)	(0.087)	(0.048)	(0.041)	(0.139)	(0.045)	(0.031)
ISales$_{t-1}$	−0.008	−0.094*	−0.091	0.080	−0.052	−0.008	−0.041	0.025	−0.067*	−0.198	−0.076	−0.083***
	(0.102)	(0.057)	(0.059)	(0.060)	(0.202)	(0.045)	(0.099)	(0.048)	(0.037)	(0.151)	(0.051)	(0.031)
Return$_{t-1}$	0.015***	0.004	0.011***	0.009***	0.015***	0.014***	−0.002	0.008**	0.015***	0.008***	0.002	0.007***
	(0.002)	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)
Sentiment$_{t-1}$	0.152***	0.165***	0.219***	0.081***	0.186***	0.146***	0.181***	0.270***	0.116***	0.157***	0.197***	0.181***
	(0.030)	(0.009)	(0.034)	(0.018)	(0.016)	(0.029)	(0.015)	(0.033)	(0.036)	(0.017)	(0.017)	(0.032)
Volume$_{t-1}$	−0.050***	0.060***	0.003	0.037***	−0.001	−0.023	0.047***	0.018	0.011	−0.013	−0.023	−0.026*
	(0.013)	(0.008)	(0.023)	(0.010)	(0.033)	(0.028)	(0.012)	(0.024)	(0.015)	(0.015)	(0.024)	(0.016)
F-statistic	18.180	7.321	37.399	9.758	9.379	17.366	7.493	53.226	9.937	26.269	53.209	23.171
Panel C: IOF$_t$												
IBuys$_{t-1}$	0.246***	0.318***	0.265***	0.281***	0.265***	0.246***	0.289***	0.217***	0.259***	0.228***	0.220***	0.241***
	(0.039)	(0.032)	(0.018)	(0.029)	(0.051)	(0.014)	(0.044)	(0.018)	(0.028)	(0.014)	(0.023)	(0.019)
ISales$_{t-1}$	−0.223***	−0.296***	−0.263***	−0.281***	−0.252***	−0.242***	−0.300***	−0.214***	−0.251***	−0.220***	−0.204***	−0.246***
	(0.034)	(0.037)	(0.018)	(0.030)	(0.055)	(0.016)	(0.055)	(0.016)	(0.026)	(0.015)	(0.023)	(0.017)
Return$_{t-1}$	0.003***	0.003***	0.006***	0.005***	0.004***	0.004***	0.001	0.002***	0.004***	0.003***	0.004***	0.005***
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)
Sentiment$_{t-1}$	−0.000	−0.001	0.000	−0.002	0.000	−0.001	−0.002***	−0.003	0.001	−0.003**	−0.001	−0.003
	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.003)	(0.001)	(0.001)	(0.001)	(0.002)
Volume$_{t-1}$	−0.002	0.001	−0.000	0.002	−0.001	−0.006**	0.002	0.002	0.001	−0.001	−0.004*	0.005
	(0.002)	(0.001)	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)	(0.002)	(0.004)	(0.002)	(0.002)	(0.003)
F-statistic	28.313	41.480	80.496	47.633	34.284	82.283	5.759	31.096	48.241	82.279	60.158	81.566

Table A7

Vector autoregressions of returns, news sentiment, and institutional trading with four lags.

The table reports estimates from panel vector autoregressions (6) with firm fixed effects. The estimates are obtained using system GMM estimation as described in Section 4.2. The dependent variables are institutional order flow (IOF), stock returns, and news sentiment. Across columns, we vary the VAR lag length L from 1 to 4. The VAR diagnostics in Panel B report the Bayesian information criterion (BIC), Akaike information criterion (AIC), mean-squared error (MSE), and residual auto- and cross-correlations. The sample contains all 1,667 NYSE stocks on all 755 trading days from 2003 through 2005. The number of observations is 1,096,514. Standard errors are reported in parentheses. Levels of significance are denoted by * (10%), ** (5%), and *** (1%).

	IOF_t	$Return_t$	$Sentiment_t$	IOF_t	$Return_t$	$Sentiment_t$	IOF_t	$Return_t$	$Sentiment_t$	IOF_t	$Return_t$	$Sentiment_t$
	$L=1$			$L=2$			$L=3$			$L=4$		
Panel A: Estimates												
IOF_{t-1}	0.259***	0.177***	0.002**	0.240***	0.169***	0.003***	0.236***	0.171***	0.002**	0.234***	0.170***	0.002**
	(0.007)	(0.016)	(0.001)	(0.008)	(0.016)	(0.001)	(0.008)	(0.017)	(0.001)	(0.008)	(0.017)	(0.001)
IOF_{t-2}				0.077***	0.032**	−0.001	0.065***	0.041**	−0.001	0.063***	0.040**	−0.001
				(0.006)	(0.016)	(0.001)	(0.006)	(0.016)	(0.001)	(0.006)	(0.016)	(0.001)
IOF_{t-3}							0.055***	−0.022	0.002**	0.047***	−0.027*	0.002*
							(0.005)	(0.015)	(0.001)	(0.005)	(0.016)	(0.001)
IOF_{t-4}										0.035***	0.017	0.000
										(0.004)	(0.016)	(0.001)
$Return_{t-1}$	0.003***	−0.005***	0.001***	0.003***	−0.005***	0.001***	0.003***	−0.005***	0.001***	0.003***	−0.005***	0.001***
	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)
$Return_{t-2}$				0.000***	0.000	0.000	0.000***	−0.002	0.000	0.000***	−0.002*	0.000
				(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)
$Return_{t-3}$							−0.001***	−0.004***	0.000	−0.001***	−0.004***	0.000
							(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)
$Return_{t-4}$										0.000	0.000	0.000
										(0.000)	(0.001)	(0.000)
$Sentiment_{t-1}$	0.001	0.093***	0.104***	0.001	0.091***	0.101***	0.001	0.093***	0.100***	0.001	0.092***	0.100***
	(0.001)	(0.014)	(0.002)	(0.001)	(0.014)	(0.002)	(0.001)	(0.014)	(0.002)	(0.001)	(0.014)	(0.002)
$Sentiment_{t-2}$				−0.001	0.018	0.032***	−0.001	0.020	0.028***	−0.001	0.021	0.028***
				(0.001)	(0.013)	(0.002)	(0.001)	(0.013)	(0.002)	(0.001)	(0.013)	(0.002)
$Sentiment_{t-3}$							−0.001	0.005	0.033***	−0.001	0.005	0.030***
							(0.001)	(0.013)	(0.002)	(0.001)	(0.013)	(0.002)
$Sentiment_{t-4}$										0.000	0.003	0.028***
										(0.001)	(0.013)	(0.002)
Panel B: Diagnostics												
AIC/BIC		−6.260/−6.260			−6.262/−6.262			−6.264/−6.264			−6.264/−6.263	
MSE	0.161	1.945	0.140	0.161	1.945	0.140	0.161	1.943	0.140	0.161	1.942	0.140
Residual auto-correlation:	−0.016***	−0.001	0.006***	−0.001	−0.002	0.007***	0.001	−0.001	0.007***	0.001	−0.002**	0.006***
Residual cross-correlation:												
IOF_t		0.056***	0.001		0.056***	0.001		0.056***	0.001		0.056***	0.001
$Return_t$			0.050***			0.050***			0.050***			0.050***

Table A8

Vector autoregressions of returns, news sentiment, and institutional trading with cross-sectional heterogeneity.

The table reports maximum likelihood estimates for vector autoregressions of returns, news sentiment, and institutional trading with random intercepts and random slopes. For each firm i and time t, the joint dynamics of $\mathbf{y}_{it} = (\mathrm{IOF}_{it}, \mathrm{Return}_{it}, \mathrm{Sentiment}_{it})'$ are

$$\mathbf{y}_{it} = \boldsymbol{\alpha}_{it} + \sum_{l=1}^{L} \boldsymbol{\lambda}_{it}^{l} \mathbf{y}_{it-l} + \boldsymbol{\varepsilon}_{it},$$

where the 3×1 vector of random intercepts is $\boldsymbol{\alpha}_{it} = \boldsymbol{\alpha}_0 + \boldsymbol{\beta} \mathbf{X}_{it} + \boldsymbol{\epsilon}_i$ with explanatory variables $\mathbf{X}_{it}$ and $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \sigma_\alpha^2)$, the 3×3 random coefficient matrices are $\boldsymbol{\lambda}_{it}^l = \boldsymbol{\lambda}_0^l + \boldsymbol{\gamma}^l \mathbf{X}_{it} + \boldsymbol{\eta}_i^l$, for $l = 1, \dots, L$ with $\boldsymbol{\eta} \sim \mathcal{N}(0, \sigma_\lambda^2)$, and $\boldsymbol{\varepsilon} \sim \mathcal{N}(0, \sigma_\varepsilon^2)$ is a 3×1 vector of innovations. We set the lag length $L = 1$. The sample contains all 1,667 NYSE stocks on all 755 trading days from 2003 through 2005. The number of observations is 1,096,514. Standard errors are reported in parentheses. Levels of significance are denoted by * (10%), ** (5%), and *** (1%).

	IOF_t	Return_t	$\mathrm{Sentiment}_t$	IOF_t	Return_t	$\mathrm{Sentiment}_t$
	(A)			(B)		
α_0	0.005***	0.104***	0.025	0.005***	0.105***	0.025
	(0.002)	(0.012)	(0.022)	(0.002)	(0.012)	(0.022)
β:						
Firm size	−0.005	−0.509***	0.042	−0.005	−0.558***	0.035
	(0.013)	(0.099)	(0.106)	(0.013)	(0.099)	(0.106)
Market-to-book	0.094	−9.684***	−7.748	0.043	−9.886***	−8.035
	(0.405)	(3.038)	(5.243)	(0.408)	(3.036)	(5.244)
Profitability	−0.030	6.471***	9.815***	−0.004	6.369***	9.721***
	(0.126)	(1.027)	(2.137)	(0.127)	(1.027)	(2.143)
Leverage	−0.005	0.056	−1.318***	−0.007	0.059	−1.316***
	(0.018)	(0.134)	(0.256)	(0.018)	(0.134)	(0.256)
No. analysts	−0.647**	−0.471	−3.060	−0.495*	−0.173	−3.071
	(0.268)	(2.047)	(3.407)	(0.269)	(2.046)	(3.405)
GIM index	−0.120	−1.106	2.022	−0.139	−1.159	2.051
	(0.135)	(0.996)	(1.716)	(0.136)	(0.996)	(1.714)
Compustat missing	0.000	−0.002	0.008	0.000	−0.003	0.008
	(0.001)	(0.008)	(0.014)	(0.001)	(0.008)	(0.014)
σ_α	0.008***	0.018***	0.104***	0.008***	0.017***	0.103***
	(0.000)	(0.006)	(0.003)	(0.000)	(0.006)	(0.003)
λ_0^l:						
IOF_{t-1}	0.256***	0.196***	0.014*	0.226***	−0.036	−0.033
	(0.002)	(0.016)	(0.008)	(0.015)	(0.103)	(0.051)
Return_{t-1}	0.003***	−0.011***	0.010***	0.001	0.002	0.006
	(0.000)	(0.002)	(0.001)	(0.001)	(0.011)	(0.004)
$\mathrm{Sentiment}_{t-1}$	0.002	0.096***	0.273***	0.001	0.238***	0.323***
	(0.001)	(0.013)	(0.008)	(0.009)	(0.090)	(0.050)
γ^l:						
$\mathrm{IOF}_{t-1} * $ Firm size				0.060	16.918***	1.107
				(0.317)	(3.414)	(0.835)
$\mathrm{IOF}_{t-1} * $ Market-to-book				2.832	47.223*	−10.671
				(3.873)	(25.266)	(10.688)
$\mathrm{IOF}_{t-1} * $ Profitability				0.173	−1.060	0.056
				(1.153)	(7.182)	(2.758)
$\mathrm{IOF}_{t-1} * $ Leverage				0.132	0.947	0.073
				(0.163)	(1.002)	(0.423)
$\mathrm{IOF}_{t-1} * $ No. analysts				7.694***	−14.992	11.233
				(2.540)	(17.743)	(8.269)
$\mathrm{IOF}_{t-1} * $ GIM index				1.266	14.706*	3.971
				(1.269)	(8.367)	(4.283)
$\mathrm{IOF}_{t-1} * $ Compustat missing				−0.000	0.099	−0.033
				(0.010)	(0.065)	(0.029)
$\mathrm{Return}_{t-1} * $ Firm size				−0.003	−0.060	−0.026
				(0.009)	(0.111)	(0.022)
$\mathrm{Return}_{t-1} * $ Market-to-book				−0.169	6.627**	0.390
				(0.235)	(2.806)	(0.921)
$\mathrm{Return}_{t-1} * $ Profitability				0.136**	−3.040***	0.563
				(0.062)	(0.746)	(0.382)
$\mathrm{Return}_{t-1} * $ Leverage				0.001	0.314***	−0.032
				(0.010)	(0.117)	(0.043)
$\mathrm{Return}_{t-1} * $ No. analysts				1.705***	−6.977***	1.208*
				(0.158)	(1.886)	(0.725)
$\mathrm{Return}_{t-1} * $ GIM index				−0.070	−1.657*	0.049
				(0.081)	(0.963)	(0.339)
$\mathrm{Return}_{t-1} * $ Compustat missing				−0.001	0.021***	−0.000
				(0.001)	(0.007)	(0.003)
$\mathrm{Sentiment}_{t-1} * $ Firm size				0.000	−0.054	−1.054***
				(0.041)	(0.389)	(0.199)

Table A8 (*continued*)

	IOF$_t$	Return$_t$	Sentiment$_t$	IOF$_t$	Return$_t$	Sentiment$_t$
	(A)			(B)		
Sentiment$_{t-1}$ * Market-to-book				−0.667	−17.625	−10.958
				(1.846)	(19.132)	(10.491)
Sentiment$_{t-1}$ * Profitability				0.907	2.491	−3.322
				(0.818)	(8.751)	(4.884)
Sentiment$_{t-1}$ * Leverage				−0.071	0.348	−0.700
				(0.096)	(0.980)	(0.559)
Sentiment$_{t-1}$ * No. analysts				−4.201***	−47.240***	−7.504
				(1.436)	(15.127)	(8.264)
Sentiment$_{t-1}$ * GIM index				0.878	−0.943	2.148
				(0.648)	(6.792)	(3.821)
Sentiment$_{t-1}$ * Compustat missing				0.001	−0.045	−0.024
				(0.005)	(0.055)	(0.031)
σ_{λ^1}:						
σ_{λ^1}(IOF)	0.074***	0.269***	0.060***	0.073***	0.269***	0.049***
	(0.002)	(0.022)	(0.016)	(0.002)	(0.022)	(0.020)
σ_{λ^1}(Return)	0.005***	0.060***	0.009***	0.005***	0.059***	0.008***
	(0.000)	(0.002)	(0.001)	(0.000)	(0.002)	(0.001)
σ_{λ^1}(Sentiment)	0.018***	0.064***	0.133***	0.018***	0.068***	0.125***
	(0.002)	(0.003)	(0.008)	(0.002)	(0.004)	(0.008)
$Corr[\eta_i^1(\text{IOF}), \eta_i^1(\text{Return})]$	0.109***	−0.114***	0.055	0.082***	−0.104***	0.087
	(0.034)	(0.042)	(0.280)	(0.034)	(0.035)	(0.339)
$Corr[\eta_i^1(\text{IOF}), \eta_i^1(\text{Sentiment})]$	−0.076	−0.230**	−0.583***	−0.063	−0.191*	−0.660**
	(0.082)	(0.117)	(0.230)	(0.084)	(0.114)	(0.322)
$Corr[\eta_i^1(\text{Return}), \eta_i^1(\text{Sentiment})]$	−0.456***	−0.550***	−0.087	−0.473***	−0.617***	−0.104
	(0.086)	(0.098)	(0.128)	(0.090)	(0.102)	(0.133)
$Corr[\eta_i^1(\text{IOF}), \epsilon_i]$	−0.035	−0.299	0.158	−0.032	−0.280	0.201
	(0.036)	(0.213)	(0.144)	(0.036)	(0.271)	(0.175)
$Corr[\eta_i^1(\text{Return}), \epsilon_i]$	0.131***	0.069	−0.055	0.134***	0.084	−0.069
	(0.038)	(0.095)	(0.084)	(0.037)	(0.101)	(0.086)
$Corr[\eta_i^1(\text{Sentiment}), \epsilon_i]$	−0.155	−0.696***	0.034	−0.160	−0.689***	0.017
	(0.097)	(0.174)	(0.063)	(0.101)	(0.178)	(0.064)
σ_ε	0.160***	1.941***	0.396***	0.160***	1.941***	0.396***
	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)
Log-likelihood	449,722	−2,284,081	−63,508	449,806	−2,283,990	−63,479

Table A9
Vector autoregressions of returns, news sentiment, and institutional trading sorted by size, number of analysts, and GIM governance index.

The table reports estimates from panel vector autoregressions (6) with firm fixed effects. The estimates are obtained using system GMM estimation as described in Section 4.2. The dependent variables are institutional order flow (IOF), stock returns, and news sentiment. We set the lag length $L = 3$. The sample contains all 1,667 NYSE stocks on all 755 trading days from 2003 through 2005. We sort all firms into terciles. The number of observations is 365,505 in each portfolio. Standard errors are reported in parentheses. Levels of significance are denoted by * (10%), ** (5%), and *** (1%).

	Firm size			No. analysts			GIM governance index		
	Small	Medium	Large	Small	Medium	Large	Small	Medium	Large
Panel A: Return_t									
IOF_{t-1}	0.150***	0.173***	0.321***	0.190***	0.133***	0.228***	0.156***	0.205***	0.200***
	(0.023)	(0.027)	(0.038)	(0.028)	(0.029)	(0.039)	(0.035)	(0.037)	(0.046)
IOF_{t-2}	0.039*	0.040	0.058	0.078***	0.030	0.094**	0.080**	0.016	0.026
	(0.022)	(0.027)	(0.038)	(0.028)	(0.027)	(0.037)	(0.033)	(0.037)	(0.046)
IOF_{t-3}	−0.051**	0.048*	−0.010	−0.063**	−0.016	0.034	0.002	−0.038	0.027
	(0.021)	(0.025)	(0.036)	(0.025)	(0.026)	(0.036)	(0.032)	(0.036)	(0.042)
$Return_{t-1}$	0.003	−0.013***	−0.012***	−0.007***	−0.011***	−0.005**	−0.009***	−0.015***	−0.022***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
$Return_{t-2}$	0.002	0.000	−0.015***	0.001	0.004*	−0.020***	−0.003	−0.003	−0.005
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)
$Return_{t-3}$	−0.004*	−0.002	−0.009***	−0.004*	−0.005**	−0.006**	0.000	−0.006**	−0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
$Sentiment_{t-1}$	0.247***	0.152***	0.061***	0.180***	0.106***	0.060***	0.101***	0.077***	0.118***
	(0.067)	(0.036)	(0.014)	(0.047)	(0.029)	(0.017)	(0.030)	(0.026)	(0.033)
$Sentiment_{t-2}$	0.092	0.012	0.019	0.076*	0.024	0.030*	−0.015	0.006	0.038
	(0.064)	(0.034)	(0.014)	(0.046)	(0.028)	(0.016)	(0.030)	(0.026)	(0.032)
$Sentiment_{t-3}$	−0.013	−0.050	0.025*	0.033	−0.008	0.012	0.002	0.012	0.008
	(0.063)	(0.034)	(0.014)	(0.045)	(0.027)	(0.016)	(0.029)	(0.025)	(0.031)
Panel B: Sentiment_t									
IOF_{t-1}	0.001	0.005**	0.002	0.001	0.005***	0.000	0.007***	0.001	0.007**
	(0.001)	(0.002)	(0.004)	(0.001)	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)
IOF_{t-2}	−0.001	−0.002	0.001	−0.001	−0.001	−0.002	−0.004**	0.002	−0.002
	(0.001)	(0.002)	(0.004)	(0.001)	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)
IOF_{t-3}	0.002*	0.003*	−0.002	0.000	0.003*	0.002	0.004*	−0.002	0.002
	(0.001)	(0.002)	(0.004)	(0.001)	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)
$Return_{t-1}$	0.000***	0.001***	0.003***	0.000	0.001***	0.003***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Return_{t-2}$	0.000*	0.000	0.000	0.000	0.000	0.000	0.000**	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Return_{t-3}$	0.000	0.000	0.000	0.000	0.000**	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Sentiment_{t-1}$	0.078***	0.073***	0.111***	0.068***	0.100***	0.109***	0.106***	0.101***	0.095***
	(0.005)	(0.004)	(0.003)	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)	(0.005)
$Sentiment_{t-2}$	0.023***	0.018***	0.032***	0.015***	0.023***	0.032***	0.036***	0.032***	0.029***
	(0.004)	(0.003)	(0.002)	(0.004)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
$Sentiment_{t-3}$	0.024***	0.013***	0.040***	0.016***	0.027***	0.038***	0.042***	0.036***	0.027***
	(0.004)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)
Panel C: IOF_t									
IOF_{t-1}	0.237***	0.234***	0.230***	0.248***	0.241***	0.224***	0.231***	0.268***	0.246***
	(0.012)	(0.004)	(0.006)	(0.015)	(0.014)	(0.007)	(0.012)	(0.010)	(0.009)
IOF_{t-2}	0.066***	0.060***	0.071***	0.057***	0.068***	0.072***	0.057***	0.060***	0.064***
	(0.009)	(0.004)	(0.004)	(0.009)	(0.012)	(0.006)	(0.011)	(0.009)	(0.006)
IOF_{t-3}	0.054***	0.061***	0.046***	0.042***	0.063***	0.050***	0.048***	0.054***	0.063***
	(0.007)	(0.003)	(0.004)	(0.007)	(0.009)	(0.005)	(0.009)	(0.008)	(0.006)
$Return_{t-1}$	0.001**	0.005***	0.005***	0.001***	0.004***	0.006***	0.004***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Return_{t-2}$	0.000	0.001***	0.000	0.000	0.001***	0.000	0.001**	0.000**	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Return_{t-3}$	0.000**	−0.001***	−0.001***	0.000	−0.001***	−0.001***	−0.001***	−0.001***	−0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Sentiment_{t-1}$	−0.001	0.005*	−0.001**	0.004	0.003	−0.003***	−0.004	0.002	0.001
	(0.011)	(0.003)	(0.001)	(0.004)	(0.004)	(0.001)	(0.003)	(0.002)	(0.002)
$Sentiment_{t-2}$	0.004	−0.005*	0.000	−0.003	−0.003	0.001	−0.001	−0.001	−0.001
	(0.010)	(0.003)	(0.001)	(0.004)	(0.004)	(0.001)	(0.003)	(0.002)	(0.003)
$Sentiment_{t-3}$	−0.010	−0.003	0.001	−0.003	0.003	−0.001	0.002	−0.001	0.001
	(0.010)	(0.003)	(0.001)	(0.004)	(0.003)	(0.001)	(0.003)	(0.002)	(0.002)

Table A10

Vector autoregressions of returns, news sentiment, and institutional trading sorted by industry.

The table reports estimates from panel vector autoregressions (6) with firm fixed effects. The estimates are obtained using system GMM estimation as described in Section 4.2. The dependent variables are institutional order flow (IOF), stock returns, and news sentiment. We set the lag length $L=3$. The sample contains all 1,667 NYSE stocks on all 755 trading days from 2003 through 2005. We sort all firms into the 12 Fama-French industries. The number of observations in each portfolio is reported in the second row of the table. Standard errors are reported in parentheses. Levels of significance are denoted by * (10%), ** (5%), and *** (1%).

	1 NoDur	2 Durbl	3 Manuf	4 Enrgy	5 Chems	6 BusEq	7 Telcm	8 Utils	9 Shops	10 Hlth	11 Money	12 Other
N	86,060	37,557	169,100	63,988	37,180	96,769	27,844	69,433	129,853	59,682	176,128	148,194

Panel A: Return$_t$

	1 NoDur	2 Durbl	3 Manuf	4 Enrgy	5 Chems	6 BusEq	7 Telcm	8 Utils	9 Shops	10 Hlth	11 Money	12 Other
IOF$_{t-1}$	0.196***	0.273***	0.178***	0.137***	0.078	0.243***	0.333**	0.491***	0.035	0.268***	0.189***	0.146***
	(0.072)	(0.102)	(0.042)	(0.046)	(0.110)	(0.058)	(0.153)	(0.106)	(0.046)	(0.067)	(0.047)	(0.037)
IOF$_{t-2}$	0.009	0.083	0.018	0.120***	0.102	−0.029	0.248	0.083	0.055	−0.033	0.013	0.060
	(0.061)	(0.113)	(0.041)	(0.046)	(0.101)	(0.058)	(0.164)	(0.100)	(0.045)	(0.069)	(0.042)	(0.041)
IOF$_{t-3}$	−0.087	−0.122	0.020	−0.111***	0.163*	0.042	−0.047	−0.031	0.027	−0.013	−0.036	−0.031
	(0.055)	(0.096)	(0.038)	(0.042)	(0.097)	(0.054)	(0.159)	(0.097)	(0.043)	(0.061)	(0.042)	(0.038)
Return$_{t-1}$	−0.019***	0.017**	−0.009***	0.007	−0.030***	0.000	−0.004	0.005	−0.001	−0.003	−0.009**	−0.006*
	(0.005)	(0.008)	(0.003)	(0.005)	(0.008)	(0.004)	(0.009)	(0.007)	(0.004)	(0.006)	(0.004)	(0.004)
Return$_{t-2}$	0.012**	0.015**	0.005	−0.037***	0.011	−0.002	−0.008	−0.018**	−0.006*	0.008	−0.002	−0.002
	(0.005)	(0.007)	(0.003)	(0.005)	(0.007)	(0.004)	(0.008)	(0.007)	(0.004)	(0.005)	(0.004)	(0.003)
Return$_{t-3}$	−0.006	0.007	0.000	−0.015***	−0.004	−0.009**	−0.016*	−0.015**	−0.004	0.003	−0.007**	0.000
	(0.005)	(0.007)	(0.003)	(0.005)	(0.007)	(0.004)	(0.008)	(0.007)	(0.003)	(0.005)	(0.003)	(0.003)
Sentiment$_{t-1}$	0.040	−0.013	0.221***	0.063	0.184***	0.066	0.117	0.079*	0.077*	0.125***	0.059**	0.095**
	(0.040)	(0.088)	(0.041)	(0.055)	(0.069)	(0.051)	(0.088)	(0.044)	(0.039)	(0.044)	(0.030)	(0.048)
Sentiment$_{t-2}$	0.060	0.067	0.026	0.096*	0.026	0.026	−0.065	0.014	−0.002	−0.022	0.074***	−0.031
	(0.039)	(0.080)	(0.042)	(0.055)	(0.067)	(0.050)	(0.084)	(0.043)	(0.038)	(0.044)	(0.028)	(0.047)
Sentiment$_{t-3}$	0.020	0.115	0.057	0.035	0.028	−0.041	−0.144*	−0.003	0.037	−0.059	0.046	−0.069
	(0.037)	(0.079)	(0.040)	(0.054)	(0.070)	(0.050)	(0.082)	(0.043)	(0.037)	(0.044)	(0.028)	(0.047)

Panel B: Sentiment$_t$

	1 NoDur	2 Durbl	3 Manuf	4 Enrgy	5 Chems	6 BusEq	7 Telcm	8 Utils	9 Shops	10 Hlth	11 Money	12 Other
IOF$_{t-1}$	0.008*	0.009*	0.000	−0.001	−0.002	0.006*	−0.010	−0.004	0.002	−0.004	−0.002	0.008***
	(0.005)	(0.005)	(0.002)	(0.002)	(0.006)	(0.003)	(0.011)	(0.010)	(0.003)	(0.003)	(0.003)	(0.002)
IOF$_{t-2}$	−0.002	−0.007	−0.003*	0.004*	−0.009	−0.004*	0.013	−0.014	−0.003	0.004	0.004*	−0.002
	(0.003)	(0.005)	(0.002)	(0.002)	(0.007)	(0.003)	(0.010)	(0.010)	(0.002)	(0.003)	(0.003)	(0.002)
IOF$_{t-3}$	0.005	−0.002	0.001	0.003*	0.003	−0.001	−0.011	0.008	0.002	0.001	0.002	0.000
	(0.004)	(0.004)	(0.002)	(0.002)	(0.005)	(0.003)	(0.011)	(0.010)	(0.002)	(0.002)	(0.003)	(0.002)
Return$_{t-1}$	0.000	0.002***	0.001***	0.001***	0.001***	0.001***	0.000	0.001	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Return$_{t-2}$	0.000	0.000	0.000	0.000	−0.001***	0.000	0.000	0.001	−0.001	0.000	0.000	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Return$_{t-3}$	0.000	0.000	0.000	0.000	−0.001*	0.000	0.000	−0.001	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Sentiment$_{t-1}$	0.091***	0.110***	0.094***	0.080***	0.086***	0.101***	0.105***	0.089***	0.082***	0.148***	0.117***	0.092***
	(0.007)	(0.012)	(0.006)	(0.007)	(0.010)	(0.007)	(0.012)	(0.007)	(0.005)	(0.008)	(0.005)	(0.006)
Sentiment$_{t-2}$	0.033***	0.016	0.032***	0.029***	0.013	0.023***	0.053***	0.020***	0.012***	0.041***	0.032***	0.032***
	(0.006)	(0.010)	(0.005)	(0.007)	(0.009)	(0.006)	(0.012)	(0.006)	(0.005)	(0.007)	(0.005)	(0.005)
Sentiment$_{t-3}$	0.026***	0.038***	0.023***	0.027***	0.030***	0.036***	0.034***	0.036***	0.015***	0.071***	0.033***	0.031***
	(0.006)	(0.010)	(0.005)	(0.007)	(0.009)	(0.006)	(0.012)	(0.006)	(0.005)	(0.007)	(0.004)	(0.005)

Table A10 (*continued*)

Panel C: IOF_t

	1 NoDur	2 Durbl	3 Manuf	4 Engy	5 Chems	6 BusEq	7 Telcm	8 Utils	9 Shops	10 Hlth	11 Money	12 Other
N	86,060	37,557	169,100	63,988	37,180	96,769	27,844	69,433	129,853	59,682	176,128	148,194
IOF_{t-1}	0.204***	0.239***	0.240***	0.293***	0.186***	0.232***	0.256***	0.204***	0.229***	0.261***	0.218***	0.229***
	(0.035)	(0.023)	(0.012)	(0.042)	(0.034)	(0.015)	(0.028)	(0.014)	(0.010)	(0.024)	(0.021)	(0.015)
IOF_{t-2}	0.052*	0.060***	0.058***	0.051***	0.080***	0.033***	0.076***	0.075***	0.083***	0.025	0.077***	0.082***
	(0.029)	(0.017)	(0.010)	(0.018)	(0.022)	(0.010)	(0.016)	(0.011)	(0.010)	(0.032)	(0.020)	(0.015)
IOF_{t-3}	0.059***	0.058***	0.051***	0.039***	0.096***	0.069***	0.040**	0.054***	0.057***	0.066***	0.044***	0.057***
	(0.022)	(0.014)	(0.009)	(0.012)	(0.037)	(0.008)	(0.018)	(0.010)	(0.008)	(0.017)	(0.017)	(0.013)
$Return_{t-1}$	0.001*	0.003***	0.003***	0.007***	0.003***	0.004***	0.001**	0.001**	0.003***	0.003***	0.002***	0.002***
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
$Return_{t-2}$	0.002***	0.001*	0.000	0.001*	0.000	0.000	−0.001	0.000	0.001*	0.001*	−0.001*	0.000
	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
$Return_{t-3}$	0.000	0.000	−0.001***	−0.001*	−0.001	−0.001***	−0.001***	0.000	−0.001**	0.000	−0.001***	0.000
	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Sentiment_{t-1}$	0.000	0.010*	0.004	−0.009*	0.001	0.003	−0.001	0.002	0.004	0.000	0.002	−0.005
	(0.007)	(0.006)	(0.003)	(0.005)	(0.008)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.005)
$Sentiment_{t-2}$	0.000	−0.004	0.000	0.001	0.004	−0.004	0.008*	−0.003	0.003	−0.001	−0.001	−0.005
	(0.005)	(0.005)	(0.004)	(0.005)	(0.006)	(0.003)	(0.005)	(0.003)	(0.003)	(0.002)	(0.004)	(0.005)
$Sentiment_{t-3}$	0.007	−0.004	−0.002	0.002	0.007	0.003	−0.003	−0.001	−0.001	−0.001	0.001	−0.012**
	(0.006)	(0.005)	(0.004)	(0.005)	(0.007)	(0.004)	(0.004)	(0.003)	(0.003)	(0.002)	(0.003)	(0.005)

References

Arellano, M., Bover, O., 1995. Another look at the instrumental variable estimation of error-components models. Journal of Econometrics 68, 29–51.

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

Auerbach, A., Gorodnichenko, Y., 2011. Measuring output responses to fiscal shocks. American Economic Journal: Public Policy 4, 1–27.

Back, K., Cao, C.H., Willard, G.A., 2000. Imperfect competition among informed traders. Journal of Finance 55, 2117–2155.

Badrinath, S., Kale, J., Noe, T., 1995. Of shepherds, sheep, and the cross-autocorrelations in equity returns. Review of Financial Studies 8, 401–430.

Baker, M., Litov, L., Wachter, J., Wurgler, J., 2010. Can mutual fund managers pick stocks? Evidence from their trades prior to earnings announcements. Journal of Financial and Quantitative Analysis 45, 1111–1131.

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

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

Boehmer, E., Wu, J., 2008. Order flow and prices. Unpublished working paper. Singapore Management University, University of Georgia.

Boudoukh, J., Feldman, R., Kogan, S., Richardson, M., 2013. Which news moves stock prices? A textual analysis. NBER Working Paper No. 18725.

Boulatov, A., Hendershott, T., Livdan, D., 2013. Informed trading and portfolio returns. Review of Economic Studies 80, 35–72.

Busse, J., Green, C., Jegadeesh, N., 2012. Buy-side trades and sell-side recommendations: interactions and information content. Journal of Financial Markets 15, 207–232.

Calza, A., Monacelli, T., Stracca, L., 2013. Housing finance and monetary policy. Journal of the European Economic Association 11, 101–122.

Campbell, J., Thompson, S., 2008. Predicting excess stock returns out of sample: Can anything beat the historical average?. Review of Financial Studies 21, 1509–1531.

Campbell, J., Ramadorai, T., Schwartz, A., 2009. Caught on tape: Institutional trading, stock returns, and earnings announcements. Journal of Financial Economics 92, 66–91.

Canova, F., 2005. The transmission of U.S. shocks to Latin America. Journal of Applied Econometrics 20, 229–251.

Canova, F., Pappa, P., 2007. Price differentials in monetary unions: the role of fiscal shocks. Economic Journal 117, 713–737.

Chan, L., Lakonishok, J., 1997. Institutional equity trading costs: NYSE versus Nasdaq. Journal of Finance 52, 713–735.

Choi, N., Sias, R.W., 2012. Why does financial strength forecast stock returns? Evidence from subsequent demand by institutional investors. Review of Financial Studies 25, 1550–1587.

Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society, Series B 39, 1–38.

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

Fang, L., Peress, J., Zhang, L., 2014. Does media coverage of stocks affect mutual funds' trading and performance? Review of Financial Studies 27, 3441–3466.

Griffin, J., Shu, T., Topaloglu, S., 2012. Examining the dark side of financial markets: Do institutions trade on information from investment bank connections?. Review of Financial Studies 25, 2155–2188.

Hamilton, J.D., 1994. Time Series Analysis. Princeton University Press, Princeton, New Jersey.

Hasbrouck, J., 1991. Measuring the information content of stock trades. Journal of Finance 46, 179–207.

Heston, S., Sinha, N., 2013. News versus sentiment: comparing textual processing approaches for predicting stock returns. Unpublished working paper. University of Maryland, Board of Governors of the Federal Reserve System.

Holmstrom, B., Tirole, J., 1993. Market liquidity and performance monitoring. Journal of Political Economy 101, 678–709.

Holtz-Eakin, D., Newey, W., Rosen, H., 1988. Estimating vector autoregressions with panel data. Econometrica 56, 1371–1395.

Infonic, 2008. Reuters NewsScope sentiment engine. White paper. Infonic Sentiment Technologies.

Irvine, P., Lipson, M., Puckett, A., 2007. Tipping. Review of Financial Studies 20, 741–768.

Jegadeesh, N., Tang, Y., 2010. Institutional trades around takeover announcements: evidence of skill and information leakage. Unpublished working paper. Emory University.

Kacperczyk, M., Seru, A., 2007. Fund manager use of public information: new evidence on managerial skills. Journal of Finance 62, 485–528.

Kaniel, R., Liu, S., Saar, G., Titman, S., 2012. Individual investor trading and return patterns around earnings announcements. Journal of Finance 67, 639–680.

Kelley, E., Tetlock, P., 2013. How wise are crowds? Insights from retail orders and stock returns. Journal of Finance 68, 1229–1265.

Kyle, A.S., 1985. Continuous auctions and insider trading. Econometrica 53, 1315–1335.

Loughran, T., McDonald, B., 2011. When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. Journal of Finance 66, 35–65.

Neuhierl, A., Scherbina, A., Schlusche, B., 2013. Market reaction to corporate press releases. Journal of Financial and Quantitative Analysis 48, 1207–1240.

Petersen, M., 2009. Estimating standard errors in finance panel data sets: comparing approaches. Review of Financial Studies 22, 435–480.

Roll, R., 1988. R^2. Journal of Finance 43, 541–566.

Savor, P., Wilson, M., 2014. Asset pricing: a tale of two days. Journal of Financial Economics 113, 171–201.

Securities Industry Association, 2007. Securities Industry Fact Book. New York 2007.

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

Sias, R.W., Starks, L., 1997. Return autocorrelation and institutional investors. Journal of Financial Economics 46, 103–131.

Sinha, N., 2012. Underreaction to news in the US stock market. Unpublished working paper. Board of Governors of the Federal Reserve System.

Tetlock, P., 2007. Giving content to investor sentiment: the role of media in the stock market. Journal of Finance 62, 1139–1168.

Tetlock, P., Saar-Techansky, M., Macskassy, S., 2008. More than words: quantifying language to measure firms' fundamentals. Journal of Finance 63, 1437–1467.

Tetlock, P., 2010. Does public financial news resolve asymmetric information? Review of Financial Studies 23, 3520–3557.