RFS Advance Access published January 24, 2015

Rumor Has It: Sensationalism in Financial Media

Kenneth R. Ahern
University of Southern California

Denis Sosyura
University of Michigan

For helpful comments, we thank two anonymous referees, David Hirshleifer (Editor), Sergey Chernenko, Zhi Da, Harry DeAngelo, Matthew Gentzkow, Tim Loughran, Chris Parsons, Joël Peress, David Solomon, Paul Tetlock, Karin Thorburn, and seminar participants at Emory University, New York University, Stockholm School of Economics, University of Alberta, University of Illinois-Urbana Champaign, University of Michigan, and University of Southern California; and conference participants at 2013 Economics of Strategy Workshop at NYU, 2013 Southern California Finance Conference, 2013 University of British Columbia Summer Finance Conference, 2013 University of Miami Behavioral Finance Conference, 2014 International Finance Conference at Universidad Católica de Chile, 2014 NBER Behavioral Finance Conference, and 2015 American Finance Association Meeting. Send correspondence to Kenneth Ahern, University of Southern California, Marshall School of Business, 3670 Trousdale Parkway, Los Angeles, CA 90089; telephone: 213-821-5583. E-mail: kenneth.ahern@marshall.usc.edu.

Rumor Has It: Sensationalism in Financial Media

Abstract

The media has an incentive to publish sensational news. We study how this incentive affects the accuracy of media coverage in the context of merger rumors. Using a novel dataset, we find that accuracy is predicted by a journalist's experience, specialized education, and industry expertise. Conversely, less accurate stories use ambiguous language and feature well-known firms with broad readership appeal. Investors do not fully account for the predictive power of these characteristics, leading to an initial target price overreaction and a subsequent reversal, consistent with limited attention. Overall, we provide novel evidence on the determinants of media accuracy and its effect on asset prices. (*JEL* G14, G34, L82)

The business press plays a key role in capital markets as a distributor of information (Tetlock 2010; Engelberg and Parsons 2011; Peress 2014). This role is not passive, however, as business newspapers actively compete for readership. To win readers' attention, newspapers have an incentive to publish sensational stories, namely attention-grabbing, speculative news with broad readership appeal. Understanding this incentive is important. Media coverage that is skewed toward speculative stories, possibly at the expense of accuracy, could distort investors' beliefs and impact asset prices. While prior research shows that the incidence of media coverage influences financial markets, there is relatively little evidence on its accuracy.

In this paper, we study accuracy in the business press in the context of merger rumors. These stories attract a broad audience because mergers have a dramatic impact on a wide range of corporate stakeholders. For employees, customers, and rivals, mergers lead to layoffs, discontinued products, and increased competition, in addition to the 15–20% abnormal return realized by target investors. At the same time, merger rumors provide a convenient setting to study accuracy in the press because we can observe ex post whether a rumor comes true.

To illustrate the trade-off between readership appeal and accuracy, consider an article that appeared on the front page of the *Seattle Times* on September 2, 1993, entitled "Could GE Buy Boeing? It's Speculation Now, But Not Entirely Far-Fetched": "A scenario by which fiercely independent Boeing succumbs to an opportunistic corporate raider has been quietly percolating in certain corners of Wall Street for the past year.. . . GE's ambitious Chairman Jack Welch, 57, has been taking steps to position GE to make a major acquisition.. . . Although he hasn't said so explicitly, Welch appears to covet Boeing." A letter to the editor that was published a few days later provides insight into how this article was received by readers. The letter states, "In my opinion, your paper chose to give this story front-page attention only for the purpose of selling newspapers. Unfortunately, judging by the fact that *The Times* newspaper box outside the gate where I work was empty when I left work (this is the first time I've noticed this occurrence), you succeeded."

This anecdote illustrates a number of interesting features of merger rumors. First, the article is designed to attract readers. Because Boeing is a major corporate presence in Seattle, a merger with GE would impact a large number of *Seattle Times* readers. Second, the article is written with provocative language that one might find in a paperback novel, such as "opportunistic raider," "Boeing succumbs," and "Welch covets." Finally, as the letter to the editor reveals, while not everyone was convinced by the article, the sensational reporting style was successful in selling newspapers. In the end, however, the rumor never materialized—GE never made a bid for Boeing.

We use merger rumors to investigate two main questions. First, which characteristics of media articles predict whether a rumor will come true? Second, do investors account for the characteristics that predict accuracy? While merger rumors allow us to address these questions in a relatively clean setting, we believe the answers can shed light on the accuracy of the business press in more general settings.

To answer these questions, we construct a novel database of merger rumors. We manually search Factiva to identify "scoop" articles that first report a merger rumor, whether they appeared in the online or print edition of a newspaper. The sample includes 501 unique merger rumors in 2000–2011, whose targets are worth 32% of the aggregate value of all acquisitions during this period. Consistent with an incentive to win readers' attention, newspapers are more likely to report rumors about newsworthy targets: large, public firms with recognizable brands and large advertising expenditures. For instance, 88% of targets in merger rumors are publicly traded, compared to 38% of targets in actual mergers. Also, 16% of rumor targets appear on league tables of the most valuable brands, compared to 1% for all merger targets.

Central to this paper is the definition of accuracy. One definition of accuracy of rumors is the literal definition. As long as a rumor is discussed in any setting, an article is literally accurate. A more relevant definition to newspaper readers is based on whether the merger materializes in the future, not whether someone is making idle speculation. Therefore, we define a rumor to be accurate if the rumor target receives an official takeover bid within one year. Using this definition, 33% of rumors in our sample are accurate. We also provide a number of robustness checks using alternative definitions of accuracy.

The accuracy of rumor articles has a significant impact on stock prices. Targets of accurate rumors earn an abnormal return of 6.9% on the rumor date, compared to 3.0% for targets of inaccurate rumors. This dichotomy implies that returns are informative about a rumor's accuracy. However, the average firm realizes a significant reversal of −1.4% over the ten days following the publication of the rumor. This finding suggests that investors overestimate the accuracy of the average rumor.

To address our first question on the determinants of accuracy, we estimate logit regressions of the likelihood that a rumor comes true based on four sets of factors: the newsworthiness of the target, characteristics of journalists, details in the text of the article, and attributes of newspapers. Because some rumors likely circulate before they are published in a newspaper, in all of our tests we control for stale information using the run-up in the target's stock price before the rumor is published.

First, we find that rumors about newsworthy targets are significantly less likely to come true. This finding suggests that newspapers may be willing to publish rumors that are less accurate if they feature large, well-known firms with broad readership appeal.

Second, characteristics of journalists significantly predict the accuracy of a rumor. Using a variety of independent sources, we hand-construct a comprehensive dataset on journalists' education, experience, and demographics. For example, we document that a third of the reporters in our sample majored in journalism in college and about half of the reporters are assigned to the New York bureau of national newspapers. A journalist is more accurate if he is older, has an undergraduate degree in journalism, and specializes in the target's industry.

Third, details in the text of the article signal a rumor's accuracy. Accurate articles are more likely to mention a specific takeover price, to discuss possible bidders, and to indicate that negotiations are in an advanced stage. Using the dictionary from Loughran and McDonald (2011), we find that an article's use of weak modal words, such as "maybe," "appears," and "conceivable," indicates that a rumor is less likely to come true.

Finally, newspaper characteristics are less important for accuracy than journalist and article characteristics. While newspaper fixed effects help to explain accuracy, a newspaper's age, circulation, form of ownership, and location do not. Overall, in answer to our first question, we find that rumors are more likely to be accurate when the target firm is less newsworthy, when journalists are more experienced, and when the article text provides specific details and uses explicit language.

To address our second question on the impact of accuracy on stock prices, we provide two sets of tests. The first set of tests shows that the magnitude of the market's immediate response to a rumor does not fully predict a rumor's accuracy. In particular, traits of newsworthy firms, characteristics of the article's text, and journalists' age, experience, and education remain significantly related to accuracy after controlling for the stock market's response to the rumor. The second set of tests shows that long-short portfolios based on the predicted accuracy of a rumor generate significant excess returns of about

1% per month.[1] Additionally, excess returns are greater when we predict a rumor's accuracy using variables that are harder to observe (e.g., journalists' age and education) compared with variables that are easier to observe (e.g., newspaper fixed effects). This suggests that the average newspaper reader overlooks relevant, but obscure, information.

We next show that the costs of arbitrage and the sophistication of investors influence asset prices following rumors. First, illiquid stocks with high idiosyncratic volatility have greater long-short portfolio returns and reverse to pre-rumor stock prices more slowly following an inaccurate rumor. Second, using daily data, we show that sophisticated short sellers correctly infer a rumor's accuracy. Finally, using daily data from a large proprietary database of institutional trades, we show that institutions are net sellers on rumors, suggesting that retail investors are likely net buyers.

These results are consistent with the theory that limited attention leads investors to overlook valuable public information (Hirshleifer and Teoh 2003; Hirshleifer, Lim, and Teoh 2011) and supporting empirical evidence (Engelberg 2008; Da, Gurun, and Warachka 2014). While this previous work has focused on the behavior of investors, we show that limited attention has important implications for the media. Because readers' attention is limited, the media competes for their attention by publishing sensational news. These news stories skew the information environment and move asset prices.

Finally, using a hand-collected sample of official merger agreements, we investigate the real effects of merger rumors. Of the 51% of agreements that refer to a rumor, the most common disclosure in the agreements is that the takeover premium is based on the target's stock price in the period before the rumor is published. We verify this effect in a large sample of merger bids. The returns of rumor targets are about 8% higher during

[1]This non-implementable long-short strategy is presented only to quantify the magnitude of the market distortion.

the run-up period and about 8% lower at the announcement. Thus, the average acquirer fully discounts the impact of a rumor on the target's stock price.

Our results have several implications. First, we challenge the view that the business press is a passive conduit of financial information. Instead, we show that the media's incentive to attract readers is associated with more speculative reporting. This underscores the distinction of media articles from corporate disclosures, which are typically more informative for large, well-known firms. Second, our results show that while the media impacts asset prices, it also introduces noise through speculative articles. Finally, we uncover important cross-sectional variation in the media's accuracy. This variation implies that the relation between information and asset prices could vary based on who is relaying the information to investors.

The central contribution of this paper is to provide new evidence on the determinants of accuracy in the business press. Previous research shows that individual investors prefer stocks with attention-grabbing news (Barber and Odean 2008; Da, Engelberg, and Gao 2011). Our findings suggest that newspapers might sacrifice accuracy in order to appeal to individual investors. This provides one explanation why investors trade stocks based on narratives in newspaper articles, despite easy access to firms' press releases and analysts' reports (Engelberg and Parsons 2011). Furthermore, because media speculation is difficult to disprove, our results help explain why media articles affect even the prices of large and widely followed stocks (Tetlock 2007). By identifying features of the text that predict accuracy, we also extend prior research on textual analysis in finance (Tetlock, Saar-Tsechansky, and Macskassy 2008; Loughran and McDonald 2011, 2014; Gurun and Butler 2012). Finally, we provide new evidence on the role of journalists in the stock market. Dougal et al. (2012) show that the identity of the authors of a popular *Wall Street Journal* column helps to predict next-day market returns. We show that accuracy varies across journalists and identify specific characteristics that help to explain this variation.

1. Data and Summary Statistics

To collect merger rumors, we use a multi-step approach. First, using a wide filter, we identify target firms named in merger rumors. After we identify the rumor target, we search for the first article to report the rumor, or as we call it, the "scoop" article.

More specifically, in the first step, we manually search the Factiva database using the following filters. First, we limit our sample dates to January 1, 2000, through December 31, 2011. Second, we search within Factiva's set of publications called "Major news and business publications: U.S." This set includes the thirty-three largest domestic newspapers. Within these bounds, we search for articles that include at least one of these words: "acquire," "acquisition," "merger," "deal," "takeover," "buyout," or "bid," and at least one of these words or phrases: "rumor," "rumour," "speculation," "said to be," or "talks." This search provides a noisy sample that we further refine by reading the articles to identify those that report merger rumors. For example, this first sample includes articles that discuss a merger and then an unrelated rumor, such as a rumor about a change in management. Once we identify a merger rumor, we extract the article text, the name of the target, the alleged bidders (if named), the media outlet, and the publication date. With very few exceptions, each scoop article mentions just one target firm.

Next, we search for the scoop article. To find the scoop, we first trace backward in time using the source of the rumor stated in the articles we have identified. When a rumor is re-reported, journalists typically cite a newspaper article that reported the story previously. In this second-pass search, we place no restriction on the newspaper's size or location. This means our sample includes foreign newspapers and small media outlets. In addition, our sample also includes online versions and blogs of print newspapers, such as Dealbook by the *New York Times*, which might publish rumors in advance of print versions. We follow the citation trail until we find an article that does not cite another media source. To verify that it is the scoop article, we search for all articles on the target firm starting one week before this potential scoop to find any previous articles on the rumor. In some cases, articles do not report a source. In these cases, we search backward in time for articles about the target firm until we find the earliest article that reports the rumor, using all sources in Factiva.

Using the scoop article date, we search for all articles that include the target's name in the following week to measure how widely a rumor is reported. From this sample, we read the articles to identify those that refer to the merger rumor. We identify separate rumors for the same target firm if a year has passed between rumors. Finally, we search through all merger bids announced between 2000 and 2012 in the SDC global merger database to identify any rumors that were followed by a formal public merger announcement.

The final sample includes 2,142 articles covering 501 rumors about 354 target firms. Targets include large, well-known firms, such as American Airlines, Alcoa, Sprint, and U.S. Steel, as well as foreign firms, such as InterContinental Hotels Group, Roche Holding, and Samsung, and private firms, such as Calvin Klein, Skype, and Groupon.

Of the 501 rumors, 167 (33.3%) were followed by a public bid for the target within one year, whether the deal was completed or not. Though we cannot know for sure whether a rumor is false, we can state that the majority of rumors do not come true.

1.1 Time-series statistics

Panel A of Table 1 presents the number of all articles, scoop articles, and public announcements by year from 2000 to 2011. There is an overall increasing pattern, with the year 2010 having the most articles and scoops (393 and 75) and the years 2004, 2009, and 2011 having the fewest. There is a positive but insignificant correlation between the number of scoop articles in a given year and the number of formal merger announcements in the SDC database (0.30, p-value = 0.34). The correlation between the percentage of rumors that emerged and the number of bids in SDC is weaker at 0.17 (p-value= 0.60). These correlations suggest that the prevalence of rumors is not closely tied to actual merger activity.

Panel B of Table 1 shows that articles follow a relatively uniform timing across calendar months. In untabulated data, we find no seasonality in total circulation for a set of prominent newspapers, consistent with uniform coverage of merger rumors by month.[2] In untabulated statistics, we find that few articles appear on Saturday or Sunday. Wednesday and Thursday are slightly more common than other weekdays for rumor articles, but overall, there is not much meaningful variation by day of the week.

1.2 Sample selection

One concern with our sample is that 501 rumors might seem relatively small compared to the total number of mergers. In particular, over the period of our sample, 2000–2011, there are 90,425 mergers in the SDC database, including public, private, and subsidiary targets. However, most of these mergers are for very small firms. The average bid price of this sample is \$228 million, or \$20.6 trillion in aggregate takeover value. In contrast,

[2]We use quarterly circulation data from the Audit Bureau of Circulations for the *Wall Street Journal* and the *New York Times* from 2005 to 2012.

the market equity of the average rumor target in our sample is about \$16 billion, or \$6.6 trillion in aggregate market equity for 414 rumors with available data. Thus, as a lower bound, our sample's market equity comprises 32% of the aggregate takeover value of the entire SDC sample.[3] This suggests that our results pertain to a large part of the universe of mergers after accounting for target size.

Second, though there are close to 100,000 mergers in SDC, it is important to consider how many mergers receive any media coverage at all. To estimate this, we search Factiva over the period 2000–2011 for all articles in the *Wall Street Journal*, either in print or online, that include any of the words "merger," "takeover," or "acquisition" in the headline. Using Factiva ticker codes, we identify the total number of unique firms that are mentioned in any of these articles. This is a conservative test because our sample also includes firms that are discussed in the article but are not involved in a merger (e.g., industry rivals, valuation comparables, etc.). We then compare these numbers to the number of unique firms in SDC.

We find that relatively few mergers receive media coverage. Using the algorithm described above, there are 8,402 unique firms mentioned in merger-related articles in the *Wall Street Journal* over 2000–2011. In comparison, there are 92,570 unique firms engaged in mergers in SDC. Thus, using these estimates, roughly 9% of mergers are covered in the media. These results show that our sample comprises a significant fraction of both the aggregate value of mergers and the sample of mergers that receive media coverage.

[3]This is a lower bound because the estimate of our sample's value excludes the takeover premium for public targets and the value of private targets, though these additional valuations are included in the SDC aggregate values.

1.3 Newsworthiness characteristics

To empirically identify the newsworthiness of firms named in merger rumors, we refer to commonly cited characteristics of newsworthiness in journalism studies: breadth, prominence, and proximity (Eadie 2009). Breadth refers to the size of the audience that would be interested in a specific firm, prominence refers to how well known is a firm, and proximity refers to how close is a firm.

We use a number of variables to measure newsworthiness. First, large public firms are more likely to interest readers because they employ more people, sell more products, and have more diverse stockholders. As a measure of firm size, we use log (book assets) from Compustat. As shown in panel A of Table 2, nearly 90% of rumor targets are publicly traded, and the average firm has book assets worth $12 billion. Second, as evident in households' stock portfolios (Frieder and Subrahmanyam 2005), firms with high brand recognition are more likely to interest a broad audience. To identify firms with recognizable brands, we use data from the marketing consultancy firms Interbrand and BrandZ, each of which publishes a list of the 100 most valuable brands in the world every year, starting in 2000 and 2006, respectively. Because these lists are so selective, we simply record a dummy variable for any target firm that appears on either list in any year from 2000 to 2011. Roughly 16% of rumor targets have valuable brands.

Additional measures of breadth and prominence are the ratio of a firm's advertising expenditures to total assets and the fraction of sales to households. Prior research shows that advertising expenditures significantly increase a firm's prominence to households and lead to greater ownership breadth, more trading, and intensified purchases by retail investors (Grullon, Kanatas, and Weston 2004; Lou 2014). The average firm in our sample has an advertising-to-assets ratio of 0.8%. To measure a firm's fraction of total sales to households, we use the fraction of sales by the target's industry that are purchased by households, according to the 1997 Input-Output tables of the Bureau of Economic Analysis. This measure identifies firms that sell more products directly to customers, compared to those that sell intermediate goods in the supply chain. About 38% of rumor targets' industry sales go to households. Our final measures of prominence are innovativeness (R&D/assets) and growth potential (Tobin's q).

Proximity implies that firms located closer to readers are more newsworthy because readers are more likely to work for such firms, buy their products, and invest in their stocks (Huberman 2001; Ivković and Weisbenner 2005). To record proximity, we use two measures of distance. In the first tests, in which we compare rumored merger targets to actual targets, we record whether a firm is domestic or foreign. In the tests of rumor targets, where we have a newspaper article for each target firm, we calculate the great-circle distance in miles between the headquarters of the firm and the newspaper. This measure also helps to account for the media slant toward local firms documented in Gurun and Butler (2012). Foreign firms account for 25% of our sample, and the average distance between a newspaper and the firm it covers is 387 miles.

1.4 Journalist characteristics

To collect biographical data for the 382 journalists who authored or coauthored any scoop article in our sample, we access a wide range of sources. We provide a detailed description of our collection methods in section 1 of the Internet Appendix and summarize the main data sources here.

First, we collect journalists' birth year and gender from the Lexis Nexis Public Records (LNPR) database. This database aggregates information on 450 million unique U.S. individuals (both alive and deceased) from sources such as drivers' licenses, property tax assessment records, and utility connection records. An older journalist could be better at assessing a rumor's accuracy than a younger journalist due to experience or better connections. This relation could also be driven by selection, in which only the more accurate journalists remain employed. Gender differences between male and female journalists may arise if female journalists have different connections to business insiders than male journalists.

We compute the average age across article coauthors and then use its logarithmic transformation in regression analysis. Panel B in Table 2 shows that the average age of journalist teams is 37 (log (age) = 3.6), and the 25th and 75th percentiles are 32 and 41 years old. For gender, we create a dummy variable equal to one if the article has any female coauthors, an outcome observed in 45% of rumors. In 17% of scoop articles, journalists are unnamed. Of the articles that report journalists' bylines, the average number of journalists per article is 1.5, with 62% of articles sole authored, 27% authored by two journalists, and 11% authored by more than two journalists.

Next, we collect data on journalists' education. We record the university attended by a journalist from biographical sketches on newspaper websites and professional networking websites. To verify a journalist's degree, year of graduation, and academic specialization, we contact the registrars of the universities attended by journalists or, if necessary, the National Student Clearinghouse, a degree-verification service provider. To verify degrees of female journalists, we use their maiden names from the LNPR database if we are unable to verify the degree under the journalist's current family name.

We record two characteristics of a journalist's education: undergraduate major and the quality of the undergraduate institution. Reporters who received more relevant academic training, such as that in journalism or business, could be better equipped to assess a rumor's accuracy and the integrity of its sources. Also, journalists who attended higher-ranked universities may have access to a more valuable alumni network, which can serve as an important channel of information transfer (e.g., Cohen, Frazzini, and Malloy 2008, 2010; Engelberg, Gao, and Parsons 2012).

We record a dummy variable equal to one if an article coauthor has an undergraduate major in one of these six academic areas: Business & Economics, Journalism, English, Political Science, History, and Other.[4] Panel B in Table 2 shows that the most common undergraduate major is Journalism (33%), followed by English (31%), History (26%), Political Science (19%), Business & Economics (10%), and Other (10%). To measure the quality of a journalist's undergraduate training, we use the university's median verbal SAT score, expressed as a percentile. Because most journalists attended liberal arts programs, the verbal score is arguably the more relevant score of quality for journalists. Table 2 shows that the journalists in our sample attended selective undergraduate programs, with a mean (median) SAT score percentile of 83.7 (87.0).

Next, we collect journalists' primary and secondary areas of professional specialization from the newspapers' biographical sketches. We conjecture that a journalist with expertise in the industry of the rumor target may be better positioned to evaluate a rumor's accuracy than a journalist specializing in another industry. In some cases, a journalist's specialization is evident from his or her professional job title (e.g., "Reporter, Automotive"), while in others, it is provided by the newspaper in the journalist's biographical sketch. We verify the reported specialization by reading samples of the journalists' articles. Then, we create a dummy variable equal to one if any of the coauthors has a

[4]Please see the Appendix for the complete list of fields that are included in each of the six categories.

primary or secondary expertise in the industry of the rumor target, using the Fama-French 17-industry classifications. In our sample, 55% of articles are written by teams with at least one journalist who is an expert in the target's industry.

Because a journalist's location may be important for access to information, we also record the geographic location of the journalist. Because many of the relevant information sources of merger rumors, such as investment bankers and stock traders, are concentrated in New York City, we create a dummy variable to identify New York–based journalists. We first identify a journalist's office location from his or her job title (e.g., "Correspondent, Atlanta Bureau") or from the newspaper's biographical sketch. Then we verify these data using journalists' residential addresses from the LNPR Database and match them to the location of newspaper bureaus. In 49.5% of articles, at least one of the article authors is stationed in New York.

Finally, we collect information on a journalist's awards, which may serve as a signal of superior skill. We consider the most prestigious journalist awards: the Pulitzer Prize, the Gerald Loeb Award, and the Society of American Business Editors and Writers (SABEW) Award. We collect information on award winners from the databases maintained by the award-bestowing organizations and record a dummy variable equal to one if any of the article's coauthors has been awarded or nominated for one of these awards. In our sample, 17.6% of articles are written by teams with an award-winning journalist.

Panel A of Table 3 presents statistics on the twelve most prolific journalists in our sample, each with at least six scoop articles. The most prolific is Dennis Berman of the *Wall Street Journal* with 24 scoops, followed by Andrew Ross Sorkin of the *New York Times* with 19 scoops, and Nikhil Deogun and Robert Frank of the *Wall Street Journal*, each with 13 scoops. In general, more prolific journalists are more accurate than the average journalist. In particular, Berman's accuracy rate is 62.5% and Sorkin's is 42.1%, above the average journalist accuracy rate of 37.6%.

1.5 Article characteristics

Using the text of the newspaper article, we record two types of information contained in the article: (i) the ambiguity of the language used in the article, and (ii) details specific to merger rumors.

First, because we focus on rumor accuracy, we study the frequency of weak modal words—a measure of an author's confidence—based on the word list for financial texts from Loughran and McDonald (2011), as updated in August 2013. This list includes twenty-six words, including "apparently," "maybe," "perhaps," and "suggests." The complete word list appears in the appendix. We predict that rumors in articles that contain a greater fraction of weak modal words are less likely to come true. When calculating the frequency of weak modal words, we are careful to avoid spurious matches. For example, the weak modal word "may" could refer to the calendar month of May, the retailer May Department Stores, or journalist's contact information (as in "the author may be reached at ...") Section 1 of the Internet Appendix explains how we address this issue.

Panel C of Table 2 shows that the mean frequency of weak modal words in merger rumors is 0.8%, noticeably higher than in annual reports (0.4%) or final IPO prospectuses (0.6%) documented in Loughran and McDonald (2011, 2013). As expected, the text of merger rumors is more speculative than that of financial disclosures.

Second, we collect details about the article that are specific to merger rumors. In particular, we collect the original source of the rumor cited in the article text. The vast majority (92%) are anonymous, with the rest made up of analysts, portfolio managers, bidder and target management, and others. We next collect the targets' comments in response to the rumor. In 46% of rumors, the target declines to comment on the rumor. In 38% of rumors, there is no mention that the newspaper attempted to contact the target for a comment. In 8% of cases, the article states that the target could not be reached, and in 4% the target denies the rumor.[5] We also record the stage of the merger talks in seven categories based on the text of the article. Panel C of Table 2 shows that most rumors are in the "Speculation" stage, accounting for 51% of the sample. The remainder is made up by "Preliminary talks" (9%), "In talks" (27%), "Preparing a bid" (4%), "Made offer" (5%), "Evaluating bids" (2%), and "For sale" (3%). We also record a number of additional variables that may signal the accuracy of a rumor. In particular, we record whether the article mentions the rumor in the headline (85%), reports the number and identity of alleged bidders (1.5 on average), and states an alleged takeover price (39%). We also count the number of rumor articles across all sources on the scoop date (1.7 on average).

1.6 Newspaper characteristics

Finally, we collect additional information about the newspapers that publish the articles in our sample. We obtain circulation and founding year from company reports and Audit Bureau of Circulation statistics. The average founding year of newspapers in our sample is 1922. The oldest newspaper in our sample is the *Times of London*, founded in 1785. The average daily circulation is 442,550 copies, and the most widely circulated newspaper is the *Wall Street Journal* with a circulation of 2,092,523 in 2011. We also identify the ultimate owner of each newspaper and record whether it is a family-run firm, which is the case for 74% of articles in the sample.

[5]In 3% of scoop articles, targets confirm the rumor. In Internet Appendix Table 1, we confirm that our results are robust to dropping these rumors.

Panel B of Table 3 presents summary statistics of the number and accuracy of articles published by the newspapers in our sample. The *Wall Street Journal* is the most prolific publisher of rumor articles, with 158 scoops, followed by *Dow Jones News Service* (67 scoops) and the *New York Times* (38 scoops). The rumors published in the *Wall Street Journal* and *Dow Jones News Service* are also more accurate than the average rumor, with accuracy rates of about 39%, compared to 33% for the average rumor. In contrast, the *Los Angeles Times* and *NYT Blogs* have accuracy rates less than 20%.

1.7 Accuracy

It is important to define accuracy in the context of merger rumors. In the literal sense, as long as any person, anywhere, with any degree of knowledge suggests to someone else that a firm is ripe for a takeover, a merger rumor published in the press is accurate. However, this is an extremely low bar for accuracy. It just implies that the journalist is not fabricating the rumor.

We define accuracy in what we believe is a more relevant way. In our setting, a rumor is accurate if it is followed by a public announcement of a proposed merger within one year, whether or not it results in a completed deal. This is the measure of ultimate interest to a newspaper's readers. The consequences of the merger, such as the premium paid to target shareholders, the change in control, and employee layoffs, are what the average reader cares about, not just that someone is making idle speculation.

As in any definition, our measure of accuracy is subjective. For instance, we could define accurate rumors as those that are followed by an official announcement within a shorter time frame. In our sample, 27.5% of rumors come true within six months and 15.8% come true within one month, compared to 33% using our twelve-month window. We could also require accurate rumors to correctly name the true bidding firm, which occurs in 15% of our sample. If we require accurate rumors to come true within one month and also correctly name the bidding firm, 8.4% of rumors in our sample are accurate.

We choose to use the more generous definition of accuracy based on the twelve-month window without the requirement to correctly name a bidder. However, in Internet Appendix Table 2, we show that our main results are unchanged using a definition of accuracy based on a rumor coming true in three, six, or nine months. We also show in Internet Appendix Table 3 that rumors that are more likely to be accurate have a significantly shorter time lag between the rumor date and the announcement of the actual merger. This helps validate our measure of accuracy.

We acknowledge that our definition of accuracy is not without limitations. An article could accurately report that two firms are in advanced merger negotiations, which then ultimately fail. This would be considered an inaccurate rumor using our definition. However, for our definition to be biased, the likelihood of deal failure would have to be systematically related to a characteristic of the merger negotiations or the firms involved that the journalist does not consider. Given that newspapers select stories to publish from a vast set of new information, it is reasonable that readers expect journalists to consider the likelihood of deal failure when they choose to publish a rumor. We provide empirical evidence to support this assumption in Section 7.3 of the paper.

2. Which Types of Rumors Are Covered by the Business Press?

We first document the characteristics of target firms in merger rumors that attract newspaper coverage. We would ideally compare firms discussed in published rumors to firms discussed in unpublished rumors. Because it is difficult to observe unpublished rumors, we use actual mergers as a benchmark for comparison. As long as the firm characteristics we document are unrelated to the likelihood that a firm is discussed in a rumor, whether published or not, using actual mergers as a comparison group is unbiased. To help ensure that this is the case, we use three samples of actual mergers as comparisons: all mergers, mergers of large public targets, and mergers of U.S. targets only. The first subsample includes all mergers in SDC from 2000 to 2011 with a deal value of at least $250 million. In the second subsample, we include publicly traded targets and set the minimum size threshold of actual merger targets such that their average value of log book assets is equivalent to that in the rumor sample. Finally, the third subsample includes only U.S. merger targets worth at least $250 million.

Table 4 presents univariate t-tests between average target characteristics in our rumor sample, compared to the three different subsamples of actual mergers. In the rumor sample, 88% of targets are publicly traded, more than double the fraction found in the universe of SDC targets (38%) or in the sample of U.S. targets (37%). We also find that the average value of book assets of rumored targets is significantly larger than that of actual merger targets. The difference between rumored targets and actual targets is even more stark for brand value. More than 15% of rumored targets have high brand values, compared to less than 1% for all mergers. Even in the size-matched sample, less than 3% of actual merger targets have high brand values. Rumored targets also spend significantly more on advertising than targets in any of the three samples of merger targets. Similarly, rumored targets sell 38% of their output to households, on average, significantly more than the 31% in all mergers and 34% in large mergers. Additionally, 75% of rumored targets are domestic firms, compared to 44% in the entire SDC sample.[6] We also find that rumored target firms spend more on R&D and have higher Tobin's q

[6]In this setting, we cannot compare actual distances between newspapers and firms because the firms in the actual merger samples do not have a newspaper associated with the merger. Because most of our rumor articles are published in U.S. newspapers, the fraction of foreign firms proxies for distance.

values than comparable large public merger targets. Analogous multivariate regressions present similar results in Internet Appendix Table 4.

These results provide consistent evidence that the financial press skews coverage toward more newsworthy firms. Rumors are more likely to be published for firms that appeal to a broader audience and have greater prominence, consistent with the theoretical models of media profit motives in Mullainathan and Shleifer (2005) and Gentzkow and Shapiro (2006).

3. How Do Merger Rumors Affect Stock Prices?

If investors can perfectly infer the likelihood that a rumor will come true, stock returns of targets on the date the rumor is published should reflect all information with no systematic overreaction. Instead, if investors incorrectly believe that rumors are more accurate than they truly are, the average target of a rumor will experience a reversal following the publication of the rumor. To test these predictions, we calculate abnormal stock returns by subtracting the daily return on the value-weighted CRSP index from the daily return of the target's stock. Cumulative abnormal returns are the time-series sum of the abnormal returns.

Table 5 and Figure 1 present cumulative abnormal returns in event time from 20 trading days before the rumor to 20 trading days after. On the date of the rumor publication (Day 0), the average target in a rumor experiences a 4.3% abnormal stock return.[7] Targets named in accurate rumors have abnormal returns of 6.9% on the rumor date, compared to 3.0% for inaccurate rumors, a highly statistically and economically significant difference. The average target experiences a substantial run-up of about 3%

[7]We use Day 0 returns throughout the paper rather than Day $-1, 0$ to be conservative. This ensures that the responses reflect the rumor article, rather than the run-up.

over the period twenty days before the rumor, with no significant difference in the run-up between accurate and inaccurate rumors. In contrast, in the 20 days following the publication of the rumor, targets in accurate rumors experience returns that are 492 basis points higher than targets in inaccurate rumors. This is driven by a significant reversal in the inaccurate rumors of -2.7%. Aggregating returns over the entire 41-day period, target firms of inaccurate rumors realize a complete reversal, where the total cumulative abnormal return is statistically indistinguishable from zero.

These results show that rumors in the press have large stock price effects. They also show that the market overreacts to the average merger rumor, suggesting that investors cannot perfectly distinguish the accuracy of merger rumors in the press. In particular, for the average rumor, there is a significant and large reversal of -1.4% over the ten days following the publication of the rumor.

4. What Predicts Accuracy in Merger Rumors?

We design a set of four tests to identify the factors that predict a rumor's accuracy and the factors that influence the stock price reaction to the rumor. In the first baseline test, we run a logit regression of the likelihood that a rumor comes true on the factors described above, controlling for year and industry fixed effects.

In the second test, we include the abnormal stock return of the target on the day the rumor is published (Day 0 return) as an explanatory variable in the logit test. This test identifies which explanatory factors are reflected in stock prices and which are not. If the day zero return reflects the likelihood that a rumor is true, then a variable that remains significantly related to the rumor's accuracy after controlling for the day zero return is not fully reflected in the market reaction to the rumor.

In the third test, we refine the target's day zero return as a control variable. In the spirit of Bhagat et al. (2005), if we ignore the time lag between the rumor and a public merger announcement, the day zero return has two components: the likelihood that the rumor will come true and the expected return of the target if the rumor does come true. Thus, the day zero return can be expressed as $r_0 = p \times r_a$, where p is the probability that the rumor comes true, and r_a is the return of the target on the day of the public announcement of the merger.[8] Rewriting this expression as $p = r_0/r_a$ isolates the component of the day zero return related to the accuracy of the rumor from the component related to the expected value of an accurate rumor.

To estimate p, we first estimate r_a from a linear regression of target announcement day returns on target size, industry, and year fixed effects in a sample of 2,555 official merger announcements of public targets over 2000 to 2011 from SDC. We use the coefficients from this model to fit estimates of $\hat{r}_a$ for each rumored target firm in our sample. The coefficient estimates are presented in Internet Appendix Table 5. We discuss alternative empirical models in Section 7.4. Using the estimate $\hat{r}_a$, we estimate $\hat{p}$.

Following this procedure, we replace the day zero return as an explanatory variable in the logit regression with $\hat{p}$ in the third test. This logit test estimates the likelihood that a rumor comes true, controlling for the component of the day zero return related to the accuracy of the rumor. This means that variables that continue to predict accuracy in this test are not fully reflected in the stock market response to the rumor. Section 5.1 also presents an analysis of long-short portfolio returns to assess whether investors fully account for the variables that predict a rumor's accuracy.

[8]The return r_a itself has two components: the probability that the deal completes and the value of the completed deal. For simplicity in our estimations and because of the noise inherent in estimating compound probabilities, we do not decompose the announcement return further.

Finally, in the fourth test, we run a regression of the day zero target abnormal returns on the same explanatory variables as in the logit tests. In addition, we control for $\hat{r}_a$, the estimated official announcement return. Because we estimate $\hat{r}_a$ using firm size, industry, and year fixed effects, these variables are omitted from this regression. Second, we do not control for estimated deal likelihood in this model because it is calculated using the day zero return. While the first three tests identify which factors predict accuracy and whether the market fully accounts for these factors, this fourth test identifies investors' beliefs about which factors influence accuracy, whether or not these factors actually predict accuracy.

In all of our tests, we also control for the staleness of the rumor. As mentioned above, newspapers are just one link in the diffusion of information from insiders to outsiders. Though our data collection process is designed to ensure that our sample correctly identifies the date and original source of the rumor among all media sources in Factiva, the rumors in our sample could circulate in other venues first. As the theoretical models of Van Bommel (2003) and Brunnermeier (2005) argue, informed traders have an incentive to leak inside information in advance of official announcements. Tetlock (2011) shows that stock returns respond less when media reports are more stale. If the staleness of the information varies across our sample firms, we could make incorrect inferences. For example, we could misinterpret a small stock price reaction to a variable that significantly explains accuracy as investor inattention, when, in fact, the price reaction is small because the information has already been incorporated in the stock price.

To control for staleness and information leakage, we use the cumulative abnormal returns of the target in the five trading days before the rumor is published. If the rumor has been widely circulated before the newspaper article is published, the pre-publication returns are expected to be higher. Internet Appendix Table 6 presents similar results using the cumulative abnormal returns over the twenty trading days before the article.

4.1 Does newsworthiness predict media accuracy?

Column 1 of Table 6 shows that the same factors that are associated with greater readership appeal are also associated with less accurate reporting. Rumors about large firms with valuable brands and greater advertising expenditures are significantly less likely to come true. These results are economically substantial. The odds ratio that a rumor comes true about a firm that does not have a valuable brand is 1.65 times as large as the odds ratio for a firm with a valuable brand. For a one-standard-deviation increase in target log(assets), the odds ratio that a rumor comes true decreases by 43%.[9]

In Column 2 of Table 6, we add the target's day zero returns. As expected, the day zero returns are positively related to accuracy. However, even after controlling for the day zero returns, the characteristics of newsworthy firms are still negatively and significantly related to rumor accuracy. In Column 3, after including the estimated deal likelihood as a control variable, the results persist. The coefficients for firm size, brand value, advertising, and industry sales to households are statistically equal across the columns.[10] This implies that including the market reaction (whether the day 0 returns in Column 2 or the estimated deal likelihood in Column 3) does not change the impact of these variables on a rumor's accuracy. Thus, these findings indicate that investors do not fully account for the incentives of newspapers to publish rumors about newsworthy firms.

Column 4 of Table 6 shows that firms with valuable brands, high Tobin's q, and low R&D expenditures experience lower returns on the rumor day. This indicates that

[9]The impact on the odds ratio is the multiplicative effect of $exp(\beta \times \Delta X)$. A one-standard-deviation increase in log(assets) generates $exp(\beta \times \Delta X) = 0.564$, equivalent to a 43% decrease in the odds ratio.

[10]In Internet Appendix Table 7, we present Wald tests of the equality of coefficients from seemingly unrelated regression models for all of the main regressions in the paper.

investors' perceptions of the rumor's accuracy are based on some characteristics, such as Tobin's q and R&D, that are not significant predictors of accuracy.

4.2 Do journalists predict rumor accuracy?

In Table 7, we run identical regressions as in Table 6, but use journalist characteristics as explanatory variables.[11] Column 1 shows that older journalists are significantly more accurate than younger journalists. Second, articles written by reporters that studied journalism in college are significantly more accurate than articles written by journalists who studied other fields, though the quality of the college, as proxied by SAT scores, is unrelated to accuracy. Third, journalists that specialize in the target's industry are more accurate. Finally, journalists based in New York City are also more accurate.

Controlling for the target's day zero stock returns in Column 2 or the estimated deal likelihood in Column 3 does not affect the results. This suggests that the market might not fully account for the predictive power of journalist age, education, and expertise. Column 4 shows that a journalist's age, education, and expertise do not affect the day zero stock returns. However, rumors written by New York–based journalists have a significantly higher stock price reaction on the day the rumor is published.

These findings are intuitive. Older journalists with more relevant experience may be better able to filter out false rumors, or they may have culled more reliable information sources than younger journalists. An undergraduate degree in journalism may equip reporters with investigative skills useful for verifying suspicious claims. The insignificant effect of SAT scores may indicate that these basic principles of journalism are taught equally at high and low ranked colleges. In contrast, inexperienced or untrained journalists may be more naïve and more easily fooled by a false rumor. Location also matters,

[11]Sample sizes vary across specifications due to data availability.

as New York–based journalists tend to be more accurate. This could occur because the best business journalists end up in New York, or because New York-based journalists have better connections to Wall Street insiders. It is reasonable that investors might not fully account for most of the journalist characteristics, given that this information is not prominently made available.

In additional tests, we measure whether investors respond to a journalist's track record. For each rumor, we calculate the average accuracy rate of the last scoop, the last two scoops, the last three scoops, and all prior scoops by a journalist. We also measure a journalist's experience by recording a variable that indicates a journalist's scoop number in chronological order during our sample period. Finally, we record the number of days from the journalist's last scoop. This may help identify whether a journalist actively publishes scoops or, instead, occasionally gets a lead. In Internet Appendix Table 8, we find no meaningful effects.[12]

Though we have identified the biographical traits of journalists that we believe are the most important for predicting accuracy, other unobserved characteristics of journalists are likely to be related to accuracy as well. In Internet Appendix Table 10, we run journalist fixed effects regressions where the dependent variables are accuracy and day zero returns. We include dummy variables only for the most prolific journalists with at least four scoop articles. Consistent with the summary statistics in Table 3, journalists Berman, Sorkin, and Sidel have positive fixed effects on the likelihood that a rumor comes true. For instance, the odds a rumor comes true are roughly six times higher if the article is written by Berman, compared to all other journalists. These results hold after controlling for the day zero return and the estimated deal likelihood.

[12]Our main results are unchanged after controlling for newspaper track records in Internet Appendix Table 9.

This evidence is consistent with the theory that limited attention may lead investors to overlook valuable public information and cause distortions in stock prices (Hirshleifer and Teoh 2003; Hirshleifer, Lim, and Teoh 2011).[13] The marginal effects for Andrew Ross Sorkin and Dennis Berman illustrate how limited attention is likely to drive these effects. Sorkin is a well known author of the best-selling book *Too Big to Fail*, which was made into a television movie for HBO. He is also known as the founder of the *New York Times* news service on mergers called *Dealbook*, which uses the masthead "DealBook with Founder Andrew Ross Sorkin." Without controlling for the day zero return, the magnitude of Sorkin's fixed effect is 1.64. However, once the day zero return is included, the fixed effect drops to 1.27, indicating that the stock returns account for Sorkin's accuracy, at least partially. Compare this to Dennis Berman, a prolific journalist with high accuracy rates, but not nearly as well-known as Sorkin. The magnitude of Berman's fixed effect is 1.79 without controlling for the day zero return. Once the day zero return is included, Berman's fixed effect remains virtually unchanged at 1.77. Rumors reported by Berman are more accurate than the average rumor, but stock prices do not reflect this additional accuracy.

4.3 Does the article text predict rumor accuracy?

In Table 8, we run identical regressions as before using article characteristics as explanatory variables. Consistent with our prediction, the first column shows a strong negative relationship between the use of weak modal words and the accuracy of a rumor. Next, an article that alleges that the firms are already engaged in merger talks is more likely to be

[13]Empirical evidence in support of limited attention has been documented in the context of financial information (Tetlock 2011; Da, Gurun, and Warachka 2014), earnings announcements (Engelberg 2008; DellaVigna and Pollet 2009; Hirshleifer, Lim, and Teoh 2009), economic shocks (Cohen and Frazzini 2008), and investment choices (Barber, Odean, and Zheng 2005; Solomon, Soltes, and Sosyura 2014).

accurate than an article that is purely speculative. An article that mentions a takeover price or lists more prospective bidders is also more likely to be accurate.[14]

All of these variables remain significant after controlling for the market response, which suggests that the language in the rumor articles is informative of the rumor's accuracy but overlooked by investors. This is consistent with Tetlock, Saar-Tsechansky, and Macskassy (2008), who find that negative words in the financial press predict lower firm earnings, but stock prices reflect this information only after a delay.

Investors do respond to some article characteristics. For example, when a rumor is covered by more newspapers, the rumor is more likely to be accurate, and this accuracy is reflected in the stock price response. In contrast, targets' day zero returns are higher when the rumor comes from an anonymous source than when there is an identified source, yet anonymity of the rumor source is unrelated to the likelihood that the rumor is accurate.

4.4 Do newspapers predict rumor accuracy?

Internet Appendix Table 13 presents the results of fixed effects regressions for newspapers on accuracy. In general, newspapers display fewer statistically significant fixed effects than journalists. This suggests that the characteristics of journalists are better predictors of accuracy than the identity of a newspaper.

Internet Appendix Table 14 shows that a newspaper's age, circulation, and ownership do not influence the accuracy of rumors. However, newspaper characteristics do influence the stock returns on the day the rumor is published. A rumor that appears in an older newspaper with a larger circulation generates greater day zero stock returns than one appearing in a younger newspaper with a lower circulation. These results are consistent

[14]We estimate identical regressions on article characteristics, in which we also include journalist fixed effects (Internet Appendix Table 11) and newspaper fixed effects (Internet Appendix Table 12). Virtually all of the results are unchanged. This means that even among the articles published by a particular newspaper or journalist, the characteristics of the text predict the accuracy of a rumor.

with the view that the media influences stock returns, even if it doesn't provide new or relevant information (Tetlock 2007, 2011).

Results in Internet Appendix Table 8 show a significant positive relationship between day 0 returns and the newspaper's accuracy rate in the last two and three scoops. This provides some evidence that investors have a greater response if an article is published by a newspaper that has recently been accurate. These results could suggest that investors are more sensitive to newspapers' reputations than journalists' reputations. This seems reasonable because it is easier for readers to keep track of the reputation of relatively few newspapers, compared to the large number of journalists. Alternatively, because the number of articles per newspaper is larger than the number per journalist, the track record measures may be more precise for newspapers than for journalists.

4.5 The correlation between newsworthiness, journalists, and articles

The above results show that characteristics of target firms, journalists, and the text of rumor articles help to predict a rumor's accuracy. These characteristics could be systematically related. For example, older journalists may use fewer weak modal words and cover less newsworthy firms. While our sample size prohibits us from including all of the numerous variables in one single regression model, Internet Appendix Tables 15, 16, and 17 present correlations between all of the variables.

The correlations reveal that rumors about large firms are more likely to state that the talks are in preliminary stages and that firms with high advertising expenses are more likely to be covered by the press in speculative articles and to use more weak modal words. These correlations provide some evidence that journalists attempt to indicate to their readers that rumors are more speculative when the target is more newsworthy. We interpret this to mean that the media follows its incentive to cover newsworthy firms even if the coverage is less likely to be accurate, but at the same time, signals that the story is potentially inaccurate.

5. The Asset Pricing Impact of Rumors and the Limits to Arbitrage

Our results show that investors ignore pertinent information and overestimate the accuracy of merger rumors, leading to distortions in the rumor target's stock price. In this section, we first quantify the economic magnitude of rumor predictability and then consider the frictions that generate abnormal returns.

5.1 Long-short portfolio returns

We form long-short portfolios of target firms based on the predicted accuracy of the rumor using the empirical models presented in Column 2 of Tables 6 (target newsworthiness), 7 (journalist characteristics), and 8 (article text characteristics). We also use journalist and newspaper fixed effects to predict accuracy. We then split the sample into firms with above-median and below-median fitted values of accuracy, based on each of the models. We put target firms with above-median accuracy in the long portfolio and firms with below-median accuracy in the short portfolio. We include the day zero returns in the prediction models, as shown in Column 2 of the tables. Thus, the first day that a firm can be placed in a portfolio is the day after the rumor is published. This makes our long-short portfolios more conservative, because we do not include the large differences in returns for accurate versus inaccurate rumors that occur on day zero. Firms are held in the portfolio for up to one year after the rumor date, and the portfolio is rebalanced daily using equal weights. If there are at least five firms in each long and short portfolio, we calculate the long-short portfolio returns. If there are fewer than five firms, we record the portfolio return as zero. This also biases our results toward zero.

Table 9 presents the long-short portfolio returns compounded to the monthly frequency over the period 2000 to 2012. Newey-West p-values are reported in parentheses. Each row presents results using a different prediction model, based on the newsworthiness of the target, the journalists' characteristics, article characteristics, and journalist and newspaper fixed effects.[15] Internet Appendix Table 18 presents the portfolio returns for the long and short positions separately.

If investors perfectly account for the characteristics of rumors, the long-run returns of the long-short portfolio should be zero. In Column 1, we find positive and significant returns using newsworthiness, journalist, and article characteristics to predict accuracy. These results suggest that the market does not fully account for all available information. Journalist and newspaper fixed effects are positive but insignificant, which suggests that the market does account for the most easily identifiable information—newspaper and journalist identities.

5.2 Limits to arbitrage

For excess returns to persist in long-run portfolio returns, there must be frictions that prevent arbitrage. We investigate this claim in four settings: long-short portfolio returns, speed of reversals for inaccurate rumors, shorting activity, and institutional trading.

5.2.1 Long-short portfolio returns and limits to arbitrage. We split our sample into above- and below-median subsamples based on two widely used measures of the costs of arbitrage: idiosyncratic volatility and Amihud illiquidity (Shleifer and Vishny 1997;

[15]These tests are not predictive regressions, and these portfolios are not an implementable trading strategy. We form portfolios using the full sample of rumors, which means we use future information that is unavailable to an investor. Moreover, they do not account for transaction costs, and the portfolio sizes are small. Instead, these tests provide an in-sample measure of the economic magnitude of the distortions associated with rumor articles.

Pontiff 2006; Amihud 2002; Hirshleifer, Teoh, and Yu 2011). Table 9 shows that the long-short portfolio returns for newsworthiness variables are driven by less liquid firms. Similarly, when a firm is illiquid, the long-short portfolio returns are significantly positive when using journalist or newspaper fixed effects to predict accuracy. However, when stocks are liquid, there are no excess returns in the long-short portfolios. Similar results hold for idiosyncratic risk. Though the statistical significance of these results is modest, it is important to recognize that these are low-power, conservative tests conducted with small samples over a year. Even in this setting, we still find evidence that price distortions are mitigated when arbitrage is less costly.

The magnitude of the differences in returns between stocks with high and low liquidity and idiosyncratic risk is comparable to prior work from more general settings. In particular, in the setting of short-selling constraints, Duan, Hu, and McLean (2010) find abnormal returns of 1.2% per month for long-short portfolios formed on idiosyncratic risk. This suggests that costly arbitrage helps to explain the significant long-short portfolio returns. Next, in Internet Appendix Table 19, we show that proportional bid-ask spreads are greater in the hard-to-arbitrage samples. In particular, liquid stocks have spreads of about 15 basis points, compared to spreads of 34 basis points for illiquid stocks. Stocks with low idiosyncratic volatility have spreads of 11 basis points, compared to spreads of 37 basis points for stocks with high idiosyncratic volatility. These results show that stocks that are more difficult to arbitrage are more costly to trade.

5.2.2 Reversals and the limits to arbitrage. Next, we expect that stock prices should take longer to reverse to pre-rumor levels following the publication of an inaccurate rumor when arbitrage is more costly. Figure 2 presents cumulative abnormal returns surrounding the publication of an inaccurate rumor based on idiosyncratic volatility (panel A) and Amihud Illiquidity (panel B). Cumulative abnormal returns that are significantly greater than zero are indicated by squares on the return lines. When cumulative abnormal returns are no longer significantly different from zero, stock prices have fully reversed.

A firm with high idiosyncratic volatility takes 21 days, on average, before its stock price fully reverses following an inaccurate rumor. A firm with low idiosyncratic volatility takes only one day. Stock prices of illiquid firms take 22 days to fully reverse after an inaccurate rumor, compared to five days for liquid firms.

5.2.3 Short-selling activity. Short-selling activity may help to understand the trading activity of informed investors following merger rumors. In particular, Boehmer, Jones, and Zhang (2008) and Kelley and Tetlock (2014) present evidence that short sellers, including retail traders, are informed. Engelberg, Reed, and Ringgenberg (2012) show that short sellers are better able to process publicly available information. In our setting, if short sellers are privately informed or just better at analyzing public information, they should be more likely to predict a rumor's accuracy. This implies that the demand for shorting targets of accurate rumors should fall, and the demand for shorting targets of inaccurate rumors should increase following the publication of a rumor.

Empirically, changes in short interest could reflect both supply and demand shifts in short selling. To isolate changes in the demand for short selling, we need to control for the supply of shares available for short selling. We exploit a recent database on equity lending from Markit, a data provider that collects short-selling activity directly from security lending desks at financial institutions (see Saffi and Sigurdsson 2011 and Engelberg, Reed, and Ringgenberg 2013 for descriptions of Markit). The advantage of the Markit database is that it includes daily data for both the number of shares lendable (a proxy for the supply of shares available for short selling) and the number of shares on loan (a proxy for shorting activity).[16] We obtain daily coverage for these two variables from July 2006 (the earliest date available) through December 2011.

Using the data from Markit, we calculate the short utilization ratio: the number of shares on loan divided by the number of shares lendable. By dividing by shares lendable, utilization accounts for changes in the supply of shares available for shorting. We calculate abnormal short utilization by subtracting the average daily short utilization over the six-month period that ends twenty days before the rumor date. To account for the delay in settlement and delivery in short sales, we record short-selling activity for day t using the data from Markit on day $t + 3$, following Geczy, Musto, and Reed (2002).

Figure 3 shows that short utilization falls substantially following the publication of accurate rumors, while utilization in inaccurate rumors stays relatively unchanged. This suggests that short sellers close their positions in targets of accurate rumors because they correctly predict that the shares will increase in value. Because the decrease in short utilization does not occur until after the rumor is published, these results are consistent with the idea that short sellers can better process public information, as in Engelberg, Reed, and Ringgenberg (2012). We acknowledge that the data do not show a sizable increase in short selling for inaccurate rumors, as might be predicted. This could reflect that verifying an accurate rumor is easier than disproving a false rumor. For instance, even a target insider may not know if another firm is planning to make a bid.

[16]The data provider reports a correlation of 90% between short interest and Markit shares loaned out. This proxy for short interest has been recently used in Jain et al. (2013) and Ljungqvist and Qian (2014).

5.2.4 Retail versus institutional trading. Next, we present evidence on institutional trading around rumors. Though some retail investors are informed, it is likely that as a class, institutional investors are more sophisticated than are retail investors.

We obtain a large proprietary database of institutional trades from 2000 to 2011 from ANcerno Ltd, a consulting firm that assists fund managers with monitoring trading costs. ANcerno reports that all of its institutional subscribers' trades must be routed via an ANcerno platform and are recorded in the database. For our purposes, the database has two advantages. First, the database allows us to observe institutional trades at a daily frequency, thus permitting clean identification around the release of merger rumors. Second, during our sample period, ANcerno covers a broad sample of 969 institutions (fund families, pension plans, and hedge fund complexes). Importantly, Puckett and Yan (2011) show that the ANcerno data are free from survivorship and backfill bias and are representative of the average institutional investor that files Form 13F.

Using the ANcerno data, we first measure the change in total CRSP volume compared to the total ANcerno volume. Compared to the daily CRSP volume over the six-month period that ends twenty days before the rumor is published, total trading volume on CRSP increases by 136% on the day the rumor is published. In contrast, the total trading volume of funds in the ANcerno data (measured as the sum of buys and sells), increases by 47%, relative to the benchmark period. These results imply that the spike in trading volume of rumor targets is driven by non-ANcerno investors. Although we do not know the identities of the non-ANcerno investors, it is reasonable to assume that they are likely retail traders, because the ANcerno data is representative of institutional investors overall.

Second, if retail investors buy shares from institutional investors after a rumor is published, we should observe a decrease in the net buying activity of institutions. Therefore, we construct the cumulative abnormal buy-sell imbalance for all institutions covered in the ANcerno database. The buy-sell imbalance is the daily number of shares of the target purchased by institutions minus the number of shares sold, normalized by the total number of shares bought or sold by institutions in the ANcerno database. We calculate the abnormal buy-sell imbalance by subtracting the average daily buy-sell imbalance over the six-month period that ends twenty days before the rumor date.

Panel A of Figure 4 shows that following the publication of a rumor, the institutional investors covered in ANcerno are net sellers: starting with the rumor, the cumulative abnormal buy-sell imbalance becomes substantially negative over the next thirty trading days. If we separate accurate and inaccurate rumors, the overall pattern is unchanged. Consistent with the "all that glitters" motivation for trading presented in Barber and Odean (2008), this pattern could reflect that individuals trade following the news story. In contrast, the institutions in our sample are not active traders on media rumors. They simply provide liquidity to individuals. Because individuals are often short-sell constrained, individuals tend to be net buyers, on average, and institutions take the other side of the trade.

To better understand what drives the buy-sell imbalance of institutional investors, we estimate changes in institutional buys and sells, relative to total volume. In particular, we calculate ANcerno buys/CRSP volume and ANcerno sells/CRSP volume. We calculate abnormal trading, as before, by subtracting the average daily trading activity in the six-month period that ends twenty days before the rumor. Panel B of Figure 4 presents the cumulative abnormal institutional buys and sells in event time. The figure shows that as a fraction of total CRSP volume, institutional investors in ANcerno place substantially fewer buy orders following the rumor. Given that Puckett and Yan (2011) show that the ANcerno funds are representative of institutions in general, it is reasonable to assume that retail investors account for much of the remaining increase in buy orders. These results are consistent with the idea that target stock returns are driven by the overreaction of relatively unsophisticated traders.

6. The Real Effects of Merger Rumors

In this section, we provide evidence on the effect of rumors on merger outcomes. First, we collect a sample of merger agreements for deals in our sample from the SEC's Electronic Data Gathering, Analysis, and Retrieval (EDGAR) system. In several cases when the merger documents are unavailable via EDGAR, we obtain merger agreements by contacting the investor relations departments of the acquirers. A typical merger agreement is between 150 and 300 pages long and contains a detailed discussion of the history of merger negotiations, the determinants of deal pricing, the opinions of financial advisors, and the disclosure of important risks, among other details. A limitation of studying merger agreements is that they are filed only for publicly announced deals and, therefore, speak only to the real effects of accurate rumors. Consequently, this approach does not allow us to evaluate the effect of rumors on the probability of deal announcement. Because the process of obtaining and processing merger agreements is labor intensive, we limit our sample to 95 agreements.

We search for the following keywords throughout the merger agreement: rumor, rumour, media, article, journal, journalist, leak, leakage, times, press, news, newspaper, speculation, speculate, and speculative. Second, we manually read each section that contains a hit, discard spurious matches, and classify relevant matches by the context of the discussion of the rumor and the effect of the rumor on the takeover process. Internet Appendix Table 20 shows the main categories of the effects of rumors and tabulates the percentage of merger agreements that discuss each effect. Section 2 of the Internet Appendix presents passages from the merger agreements to illustrate the language used in each category.

First, though all 95 of the mergers were preceded by a rumor in the press, approximately one-half of the firms (50.5%) discuss rumors in their merger agreements. Such discussions appear most frequently in the sections that present the narrative history of merger negotiations (68.8%), the opinion of financial advisors (41.7%), and the recommendations of the boards of directors (25.0%). Less frequently, media rumors are mentioned in the market value analysis (4.2%), executive summary (4.2%), or appendixes (2.1%).[17]

Second, among the firms that discuss merger rumors, almost two-thirds (62.5%) acknowledge the effect of the rumor on the market price of the target and state that the premium paid for the target is based on the target's price in a period that precedes the rumor. These discussions suggest that the negotiating parties actively discount the run-up in the target's stock price driven by the rumor when determining the premium. In an additional 12.5% of the cases, the merger agreement acknowledges the effect of the rumor on the target's stock price, but does not explicitly state how it affects the premium. A sizable fraction of merger agreements explicitly mention that the rumor led the involved parties to call a meeting to discuss the rumor (20.8%) and to expedite the negotiation process (4.2%). Finally, approximately 14.6% of agreements mention that a rumor triggered the negotiating parties to issue a comment on the rumor.

The evidence from merger agreements shows that the most frequently cited effect of the rumor is that the negotiating parties recognize the change in the target's stock price attributable to the rumor and fully discount such an effect in the determination of the takeover premium. We next verify whether this is true for the average merger.

[17]Percentages do not sum to 100% because some discussions appear in multiple sections of the agreement.

Schwert (1996) shows that takeover premiums include two components: the run-up in the target's stock price before the announcement of a merger and the markup from the announcement to the close of the merger. If the bidder believes the run-up simply reflects the anticipation of the upcoming merger bid, it would revise its takeover price down accordingly. Schwert finds the opposite: the run-up is an added cost to the bidder, and there is no trade-off between the run-up and the markup.

To test the effect of rumors on takeover premiums, we include the full sample of official merger bids in SDC over the period 2000–2011 for public targets and record a dummy variable equal to one if the deal was preceded by a merger rumor identified in our main sample. Following Schwert (1996), we calculate the target's cumulative abnormal stock return over the period from 42 trading days before the public announcement of the merger until one day before the announcement. The second period is the period from the day of the public announcement to five days after the announcement.[18] The total premium is calculated as the target's cumulative abnormal returns from 42 days before the announcement to five days after it. We then regress the total premium on the rumor dummy variable, plus a host of factors that might influence target returns, including target size, industry fixed effects, and deal characteristics.

Table 10 presents the results from these regressions. Consistent with our prior findings, rumors increase target returns in the run-up period by about 8 percentage points, on average. This is true after accounting for variables that could affect the accuracy of the rumor, such as brand value and size, as well as deal characteristics, such as payment method and the use of takeover defenses. The second set of regressions shows that rumors have a strong and statistically significant negative effect on target returns at

[18]Schwert (1996) extends this period for a longer duration, but the vast majority of the returns occur within the first few days of the announcement.

the announcement of about 8 percentage points. Thus the markup for rumored deals is substantially reduced. Finally, the third set of regressions shows that rumors have no significant effect on the total takeover premium. The marginal effect of the rumor variable is insignificant and economically minuscule.

These results show that rumors do not contribute to the premium paid in mergers. Consistent with the textual analysis of merger agreements, insiders appear to correctly attribute the additional stock returns caused by the rumor and adjust takeover prices downward accordingly.

7. Robustness

7.1 Do bidders' characteristics attract media coverage?

Internet Appendix Table 21 shows that bidders are also newsworthy firms. They are large, public firms with valuable brands and high advertising expenses. However, Internet Appendix Table 22 shows that the magnitudes of the logit coefficients on the newsworthiness variables for targets are all statistically greater predictors of media coverage than the corresponding coefficients for bidders. These results suggest that target characteristics have a stronger relationship with media coverage than do bidder characteristics.

Next, Internet Appendix Table 23 presents abnormal returns of the median bidder surrounding the rumor publication date. In contrast to the target's average return of 4.3% on day 0, the average returns to bidders are −0.1%, a result that is both statistically and economically insignificant. This implies that investors do not react to news about the bidder, but they do react strongly to news about the target. Thus, the media's impact on target stock returns is much greater than its impact on bidder stock returns.

7.2 Do extreme returns cause media coverage?

Though we argue that the media covers large, public firms, with high brand values because such firms interest the average reader, an alternative explanation is that journalists investigate firms that have abnormal returns or trading patterns, and mention the most extreme cases. To test this hypothesis, we regress run-up returns on newsworthiness variables in Internet Appendix Table 24. If journalists prefer to cover stocks with extreme run-ups, rather than newsworthy stocks, we would expect to find a positive relationship between newsworthiness and stock returns. Our empirical results do not support this view. Run-ups are negatively related to firm size and unrelated to all other newsworthiness characteristics except R&D. This suggests that newsworthiness is unlikely to be confounded by extreme returns that lead to the publication of a merger rumor.

7.3 Likelihood of withdrawals

One concern with our measure of accuracy is that rumors about merger negotiations that do not advance to a public bid could be classified as inaccurate, even if there were actual merger talks happening. This could confound our tests if more newsworthy firms are also more likely to engage in negotiations that ultimately fail. A direct test of this alternative explanation requires a sample of all rumors, both published and unpublished, and their outcomes. Because we cannot observe such a sample, we use a similar setting where we can identify negotiation failures: withdrawals of public merger bids.

We select a sample of official merger announcements that are chosen to match the characteristics of our rumor sample. Because our test relies on variation in newsworthiness variables, we match our sample based on other variables. In particular, we follow prior literature on the determinants of deal withdrawals and match on year, industry, firm, and deal characteristics that are likely to influence the likelihood of a withdrawn bid, including controls for form of payment, tender offers, leveraged buyouts, cross-border deals, and the use of target takeover defenses (Baker and Savaşoglu 2002; Golubov, Petmezas, and Travlos 2012). For each rumor target, we select three non-rumored deals from SDC using Mahalanobis distance of these matching variables to select the best matches.

T-tests in Internet Appendix Table 25 show that there is no significant difference between the two samples across almost all matching variables. Internet Appendix Table 26 presents logit regressions of the likelihood of a withdrawn bid using the matched sample. We find no relationship between newsworthiness and the likelihood of a withdrawn bid. This provides evidence that more newsworthy firms are not more likely to engage in negotiations that ultimately fail.

7.4 Alternative prediction models

In the Internet Appendix, we provide a wide range of alternative models to those presented in Section 4. First, Internet Appendix Table 5 presents four different regression models to explain target returns at the merger announcement. The model in Column 1 is our base model and includes target industry, size, and year fixed effects. The other columns include additional target and bidder variables to explain returns, including return on assets (ROA), Tobin's q, leverage, capital expenditures/assets, intangibles/assets, and profitability. The adjusted R^2 increases with more variables, but the sample size shrinks to one-third of the original sample. In general, the bidder variables remain largely insignificant, while target q, leverage, intangibles, and profitability significantly help to explain target returns.

Next, we construct a number of alternative estimates of $\hat{p}$, the estimated likelihood of the deal, using the new estimates of r_a. Internet Appendix Table 27 presents logit regressions, identical to specification 3 of Tables 6, 7, and 8, except that we use a variety of alternative measures of $\hat{p}$ to account for the market's estimated deal likelihood. These include winsorizing $\hat{r_a}$ to be non-negative, measuring day 0 returns using the six-day returns over the period -5 to day zero, log-transforming $\hat{p}$ by $\ln\left(\frac{\hat{p}}{1-\hat{p}}\right)$, and using estimates of $\hat{r_a}$ that include bidder industry, size, and additional firm characteristics. Across all sets of tests, our main results remain almost identical.

In Internet Appendix Table 28, we run the same regressions as in Column 4 of Tables 6, 7, and 8, where we include alternative estimates of r_a using the three additional models of announcement returns in Columns 2–4 of Internet Appendix Table 5. Across all specifications, our main results are virtually unchanged.

7.5 Legal aspects of merger rumors

In Section 3 of the Internet Appendix, we provide a detailed discussion of the legal environment surrounding the publication of rumors. In summary, it is highly unlikely that a journalist or newspaper would be charged with securities fraud for publishing a merger rumor. To prove its case, the SEC would have to provide evidence that a journalist intentionally spread a false rumor in order to realize a monetary gain. This is a very high hurdle, and we are not aware of any case in which a journalist has been charged with market manipulation through spreading false rumors.

8. Conclusion

This paper shows that media coverage of merger rumors is biased toward newsworthy firms that appeal to a broad audience. At the same time, rumors about more newsworthy firms are substantially less likely to come true. However, stock returns do not reflect the reduced accuracy related to newsworthiness.

We also provide new evidence that the biographical traits of journalists are strong predictors of accurate reporting. Older reporters who received degrees in journalism and specialize in the rumor target's industry are significantly more accurate. In addition, the specific language used in the text of a media article helps to predict whether the rumor is accurate. For example, a discussion of a specific takeover price, the disclosure of potential bidders, and the use of weak modal words that indicate uncertainty provide important signals of a rumor's accuracy. Nevertheless, the returns of portfolios sorted by rumor accuracy show that investors do not fully account for important predictors of future stock returns.

We believe our results have important implications for the role of the financial media in the stock market that extend beyond merger rumors. In particular, our results suggest that the media selectively provides more information about large, public firms with wide readership appeal, but this information is likely to be less accurate. The same bias could influence media coverage of earnings reports, executive turnover, financial distress, or any other significant corporate event that attracts media attention.

Appendix

Variable Definitions

Newsworthiness variables

Target book assets	Total book assets, as reported in Compustat.
Public target	Dummy variable equal to one if the rumor target is publicly traded at the time of the rumor.
Valuable brand	Dummy variable equal to one if the target firm was listed in the top 100 most valuable brands by the Interbrand or Brandz data in any year from 2000 to 2011.
Advertising/Assets	Advertising expenses/Total book assets, as reported in Compustat.
Industry sales to households	The fraction of the target industry's sales that are purchased by households. Data are from the 1997 Bureau of Economic Analysis Detailed-level Input-Output tables.
Tobin's q	(Total assets − common equity + market equity)/Total assets. Data from CRSP and Compustat.
R&D/Assets	R&D/Total book assets, as reported in Compustat.
Distance	Great circle distance in miles between the headquarters of the newspaper that published the scoop article and the target firm.
Foreign target	Dummy variable equal to one if the rumor target is headquartered outside of the United States.

Journalist variables

Age	The average age (in years) of all journalists listed as authors of a scoop article.
Undergraduate major	Dummy variable equal to one if an article is written by a journalist who graduated with a major in one of the following categories:
Business & Economics	Business, economics, finance, and management
Journalism	Broadcasting, communication, journalism, mass media, and media studies
English	Creative nonfiction, English, literature, literary studies, and screenwriting
Political Science	Government, international affairs, international relations, law, politics, political science, public policy, and public relations

History	Ancient history, American studies, art history, Asian history, Chinese history, classics, history, and modern history
Other	Animal science, anthropology, biology, biopsychology, criminal justice, East Asian languages, East Asian studies, electrical engineering, environmental biology, film, general studies, Germanic studies, human development, liberal arts, mathematics, philosophy, psychology, religion, Russian studies, sociology, teaching, urban affairs, veterinary medicine, and zoology.
College SAT percentile	The average verbal SAT percentile of the undergraduate institutions of all journalists listed as authors of a scoop article.
Expert in target industry	Dummy variable equal to one if any journalist who authored an article is an expert in the same industry as the primary industry of the rumor target, using Fama-French 17 industry codes.
New York–based	Dummy variable equal to one if at least one of the authors of an article is based in New York City.
Award winner	Dummy variable equal to one if at least one of the authors of an article has been nominated for or received the Pulitzer Prize in Journalism, the Gerald Loeb Award, or the Society of American Business Editors and Writers (SABEW) award.
Gender	Dummy variable equal to one if an article has at least one female coauthor.
Article variables	
Weak modal words	The fraction of weak modal words in the text of an article. Weak modal words are defined in Loughran and McDonald (2011) and include the following: *apparently, appeared, appearing, appears, conceivable, could, depend, depended, depending, depends, may, maybe, might, nearly, occasionally, perhaps, possible, possibly, seldom, seldomly, sometimes, somewhat, suggest, suggests, uncertain,* and *uncertainly*.
Anonymous source	Dummy variable equal to one if an article does not identify a specific source of the rumor.
Target comment	Categorical variable that records the target firm's response to the rumor, according to the text of the newspaper article: No comment, Has conversations from time to time, Confirmed rumor, Denied rumor, Couldn't be reached, or Wasn't asked.

Merger stage	Categorical variable that records the stage of the rumored talks, according to the text of the newspaper article: Speculation, Preliminary talks, In talks, Made offer, Preparing a bid, For sale, or Evaluating bids
Articles on scoop date (#)	The total number of articles reporting the rumor published on the same date as the scoop article.
Rumor in headline	Dummy variable equal to one if the rumor article refers to the rumor in the headline of the article.
Number of bidders mentioned	The number of firms mentioned in the text of the article as potential bidders.
Price mentioned	Dummy variable equal to one if a specific takeover price is mentioned in the text of the article.
Newspaper variables	
Family-run media company	Dummy variable equal to one if a newspaper is owned by a family-run firm.
Newspaper age	The age of the newspaper in years from its original founding date to the date of article publication.
Newspaper circulation	The total daily circulation of the newspaper, as recorded in the Audit Bureau of Circulation reports.
Other Control Variables	
Amihud illiquidity	Average $\left(\frac{\lvert r_t \rvert}{\text{Volume}_t}\right)$ as calculated by Amihud (2002).
Completed	Dummy variable equal to one if a merger bid is successfully completed, as reported in SDC.
Cross-border	Dummy variable equal to one if a merger bid is a cross-border bid, as reported in SDC.
Day 0 return	The abnormal stock return of the target firm on the day the scoop article is published. Abnormal returns are calculated as the firm's return minus the CRSP value-weighted index return.
Estimated announcement return	The fitted value of the expected announcement return of the target of an actual merger announcement. Fitted values are based on the coefficients in Internet Appendix Table 5.
Estimated deal likelihood	Day 0 return/Estimated announcement return
Idiosyncratic volatility	The standard deviation of the residual in the time-series regression of a firm's excess daily stock returns over the risk-free rate on the Fama-French factors, where observations are over the one-year period that ends 20 days before the publication of a rumor. We require at least 100 observations of daily returns to calculate idiosyncratic volatility.

Industry fixed effects	Dummy variables for the target firm's primary Fama-French 17 industry code.
Institutional buy-sell imbalance	The daily number of shares bought minus the number of shares sold divided by the sum of the number of shares bought and sold, using all funds in the ANcerno database.
Leveraged buyout	Dummy variable equal to one if a merger bid is classified as a leveraged buyout, as reported in SDC.
Majority cash	Dummy variable equal to one if a merger bid uses cash as the majority form of payment, as reported in SDC.
Returns$_{(-5,-1)}$	The cumulative abnormal stock returns over the period from five days to one day before the scoop article is published. Abnormal returns are calculated as the firm's return minus the CRSP value-weighted index return. Cumulative returns are the sum over the five days of the abnormal returns.
Target market equity	The target stock price times the number of shares outstanding two days before the announcement of the merger.
Target takeover defenses	Dummy variable equal to one if a target employed any defensive antitakeover provisions following an unsolicited merger bid, as reported in SDC.
Tender offer	Dummy variable equal to one if a merger bid is a tender offer, as reported in SDC.
Utilization ratio	The ratio of a firm's number of shares on loan to the number of shares lendable, using daily data from Markit.
Year fixed effects	Dummy variables for the year the scoop article is published.

References

Amihud, Y. 2002. Illiquidity and stock returns: Cross-section and time series effects. *Journal of Financial Markets* 5:31–56.

Baker, M., and S. Savaşoglu. 2002. Limited arbitrage in mergers and acquisitions. *Journal of Financial Economics* 64:91–115.

Barber, B., and T. Odean. 2008. All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies* 21:785–818.

Barber, B. M., T. Odean, and L. Zheng. 2005. Out of sight, out of mind: The effects of expenses on mutual fund flows. *Journal of Business* 78:2095–119.

Bhagat, S., M. Dong, D. Hirshleifer, and R. Noah. 2005. Do tender offers create value? New methods and evidence. *Journal of Financial Economics* 76:3–60.

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

Brunnermeier, M. K. 2005. Information leakage and market efficiency. *Review of Financial Studies* 18:417–57.

Cohen, L., and A. Frazzini. 2008. Economic links and predictable returns. *Journal of Finance* 63:1977–2011.

Cohen, L., A. Frazzini, and C. Malloy. 2008. The small world of investing: Board connections and mutual fund returns. *Journal of Political Economy* 116:951–79.

———. 2010. Sell side school ties. *Journal of Finance* 65:1409–37.

Da, Z., J. Engelberg, and P. Gao. 2011. In search of attention. *Journal of Finance* 66:1461–99.

Da, Z., U. G. Gurun, and M. Warachka. 2014. Frog in the pan: Continuous information and momentum. *Review of Financial Studies* 27:2171–218.

DellaVigna, S., and J. Pollet. 2009. Investor inattention and Friday earnings announcements. *Journal of Finance* 64:709–49.

Dougal, C., J. Engelberg, D. Garcia, and C. A. Parsons. 2012. Journalists and the stock market. *Review of Financial Studies* 25:639–79.

Duan, Y., G. Hu, and R. D. McLean. 2010. Costly arbitrage and idiosyncratic risk: evidence from short sellers. *Journal of Financial Intermediation* 19:564–79.

Eadie, W. F., ed. 2009. *Twenty-first century communication*. Thousand Oaks, CA: Sage.

Engelberg, J. 2008. Costly information processing: Evidence from earnings announcements. Working Paper, University of California–San Diego.

Engelberg, J., P. Gao, and C. A. Parsons. 2012. Friends with money. *Journal of Financial Economics* 103:169–88.

Engelberg, J., and C. A. Parsons. 2011. The causal impact of media in financial markets. *Journal of Finance* 66:67–97.

Engelberg, J. E., A. V. Reed, and M. C. Ringgenberg. 2012. How are shorts informed?: Short sellers, news, and information processing. *Journal of Financial Economics* 105:260–78.

———. 2013. Short selling risk. Working Paper, University of California–San Diego, University of North Carolina–Chapel Hill, and Washington University.

Frieder, L., and A. Subrahmanyam. 2005. Brand perceptions and the market for common stock. *Journal of Financial and Quantitative Analysis* 40:57–85.

Geczy, C. C., D. K. Musto, and A. V. Reed. 2002. Stocks are special too: An analysis of the equity lending market. *Journal of Financial Economics* 66:241–69.

Gentzkow, M., and J. M. Shapiro. 2006. Media bias and reputation. *Journal of Political Economy* 114:280–316.

Golubov, A., D. Petmezas, and N. G. Travlos. 2012. When it pays to pay your investment banker: New evidence on the role of financial advisors in M&As. *Journal of Finance* 67:271–311.

Grullon, G., G. Kanatas, and J. P. Weston. 2004. Advertising, breadth of ownership, and liquidity. *Review of Financial Studies* 17:439–61.

Gurun, U. G., and A. W. Butler. 2012. Don't believe the hype: Local media slant, local advertising, and firm value. *Journal of Finance* 67:561–98.

Hirshleifer, D., S. S. Lim, and S. H. Teoh. 2009. Driven to distraction: Extraneous events and underreaction to earnings news. *Journal of Finance* 64:2289–325.

———. 2011. Limited investor attention and stock market misreactions to accounting information. *Review of Asset Pricing Studies* 1:35–73.

Hirshleifer, D., and S. H. Teoh. 2003. Limited attention, information disclosure, and financial reporting. *Journal of Accounting and Economics* 36:337–86.

Hirshleifer, D., S. H. Teoh, and J. J. Yu. 2011. Short arbitrage, return asymmetry, and the accrual anomaly. *Review of Financial Studies* 24:2429–61.

Huberman, G. 2001. Familiarity breeds investment. *Review of Financial Studies* 14:659–80.

Ivković, Z., and S. Weisbenner. 2005. Local does as local is: Information content of the geography of individual investors' common stock investments. *Journal of Finance* 60:267–306.

Jain, A., P. K. Jain, T. H. McInish, and M. McKenzie. 2013. Worldwide reach of short selling regulations. *Journal of Financial Economics* 109:177–97.

Kelley, E. K., and P. C. Tetlock. 2014. Retail short selling and stock prices. Working Paper, University of Arizona and Columbia University.

Ljungqvist, A., and W. Qian. 2014. How constraining are limits to arbitrage? Evidence from a recent financial innovation. Working Paper, New York University.

Lou, D. 2014. Maximizing short-term stock prices through advertising. Working Paper, London School of Economics.

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

———. 2013. IPO first-day returns, offer price revisions, volatility, and Form S-1 language. *Journal of Financial Economics* 109:307–26.

———. 2014. Measuring readability in financial disclosures. *Journal of Finance* 69:1643–71.

Mullainathan, S., and A. Shleifer. 2005. The market for news. *American Economic Review* 95:1031–53.

Newey, W. K., and K. D. West. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55:703–8.

———. 1994. Automatic lag selection in covariance matrix estimation. *Review of Economic Studies* 61:631–53.

Peress, J. 2014. The media and the diffusion of information in financial markets: Evidence from newspaper strikes. *Journal of Finance* 69:2007–43.

Pontiff, J. 2006. Costly arbitrage and the myth of idiosyncratic risk. *Journal of Accounting and Economics* 42:35–52.

Puckett, A., and X. Yan. 2011. The interim trading skills of institutional investors. *Journal of Finance* 66:601–33.

Saffi, P. A., and K. Sigurdsson. 2011. Price efficiency and short selling. *Review of Financial Studies* 24:821–52.

Schwert, G. W. 1996. Markup pricing in mergers and acquisitions. *Journal of Financial Economics* 41:153–92.

Shleifer, A., and R. W. Vishny. 1997. The limits of arbitrage. *Journal of Finance* 52:35–55.

Solomon, D. H., E. F. Soltes, and D. Sosyura. 2014. Winners in the spotlight: Media coverage of fund holdings as a driver of flows. *Journal of Financial Economics* 113:53–72.

Tetlock, P. C. 2007. Giving content to investor sentiment: The role of media in the stock market. *Journal of Finance* 62:1139–68.

———. 2010. Does public financial news resolve asymmetric information? *Review of Financial Studies* 23:3250–557.

———. 2011. All the news that's fit to reprint: Do investors react to stale information? *Review of Financial Studies* 24:1481–512.

Tetlock, P. C., M. Saar-Tsechansky, and S. Macskassy. 2008. More than words: Quantifying language to measure firms' fundamentals. *Journal of Finance* 63:1437–67.

Van Bommel, J. 2003. Rumors. *Journal of Finance* 58:1499–519.

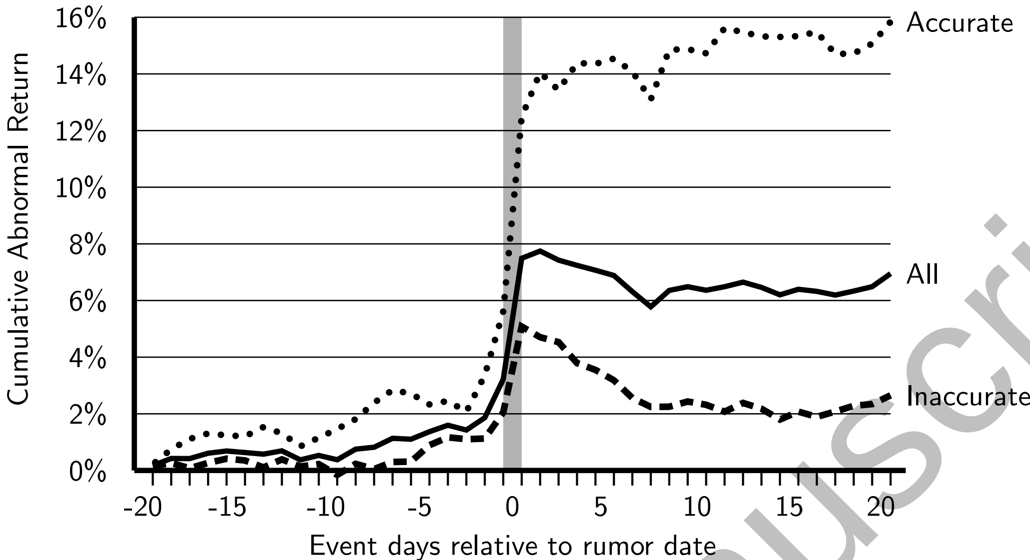
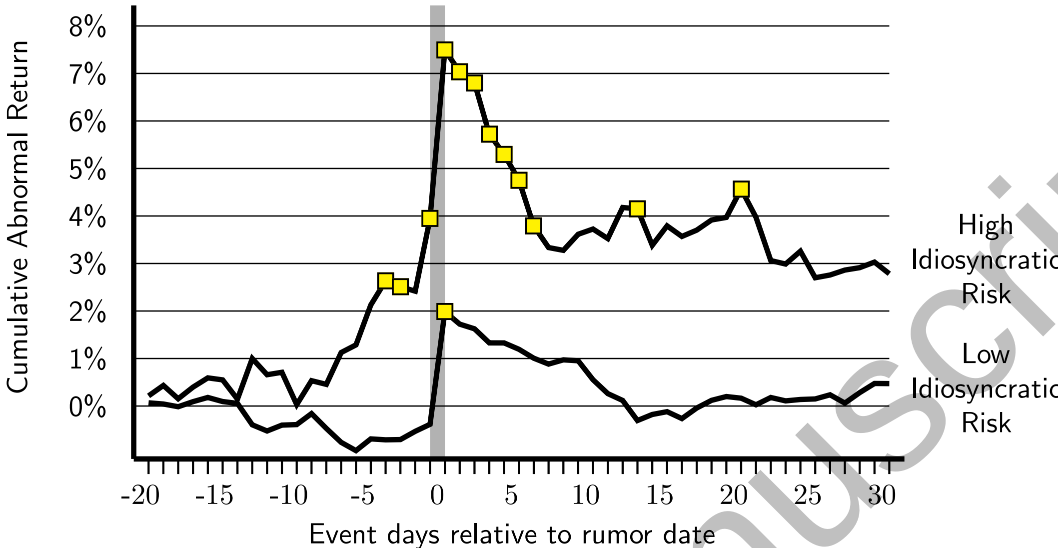
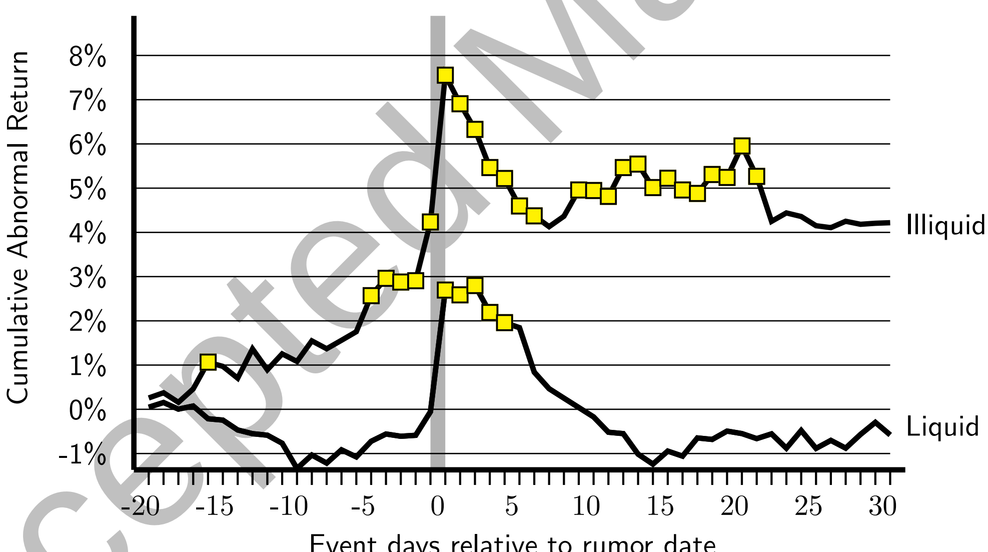


Figure 1
Abnormal returns of accurate and inaccurate merger rumors
This figure presents the time series of cumulative abnormal returns of merger rumor targets for three time series relative to the date of the first publication of a merger rumor. There are 414 rumors with stock price data over 2000–2011. Cumulative abnormal returns are the cumulative sum of daily abnormal returns, calculated as the firm's daily return minus the value-weighted CRSP index. The solid line represents the average of all target firms. The dotted line represents the subsample of merger rumor targets where a public merger announcement was made within one year of the rumor date. The dashed line represents the subsample of merger rumor targets where no public merger announcement was made in the following year.



(A) Idiosyncratic Volatility



(B) Amihud Illiquidity

Figure 2
Reversals following inaccurate rumors by limits to arbitrage
This figure presents the time series relative to the date of the first publication of a merger rumor of cumulative abnormal returns of targets in rumors that do not come true within one year of the publication of the rumor. Cumulative abnormal returns are the cumulative sum of daily abnormal returns, calculated as the firm's daily return minus the value-weighted CRSP index. The squares indicate cumulative returns that are statistically significant at the 10% level. Panel A presents abnormal returns of targets with above- or below-median idiosyncratic volatility. Panel B presents abnormal returns for targets with above or below median Amihud illiquidity.

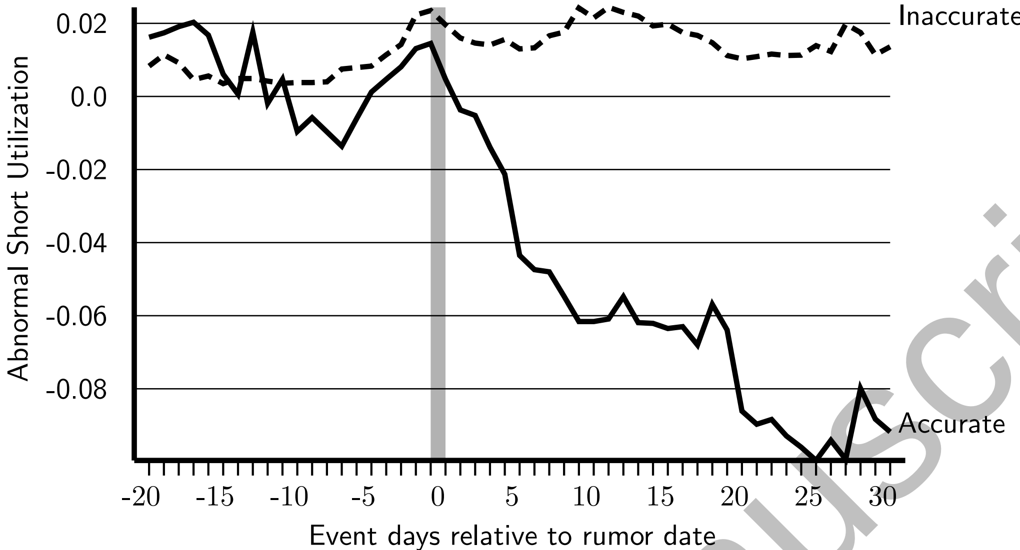
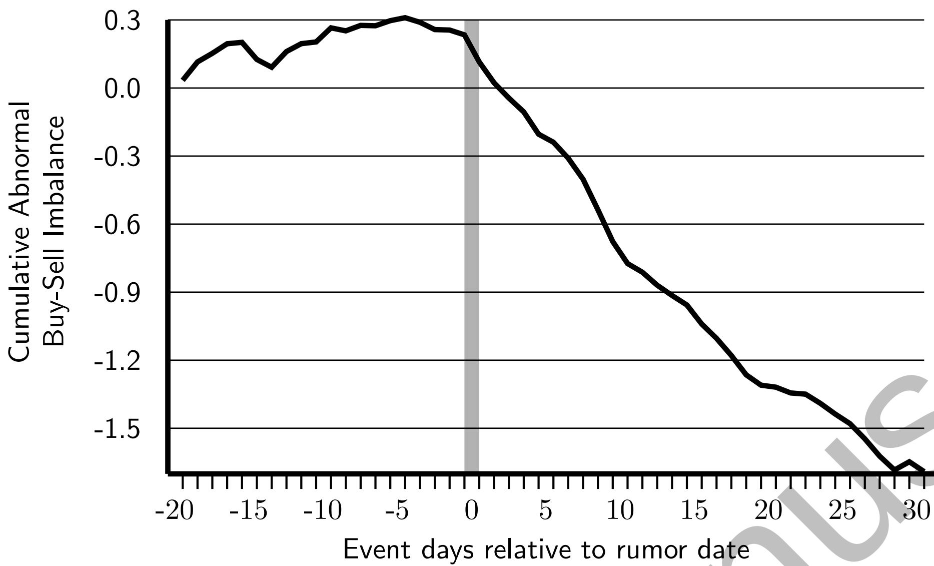
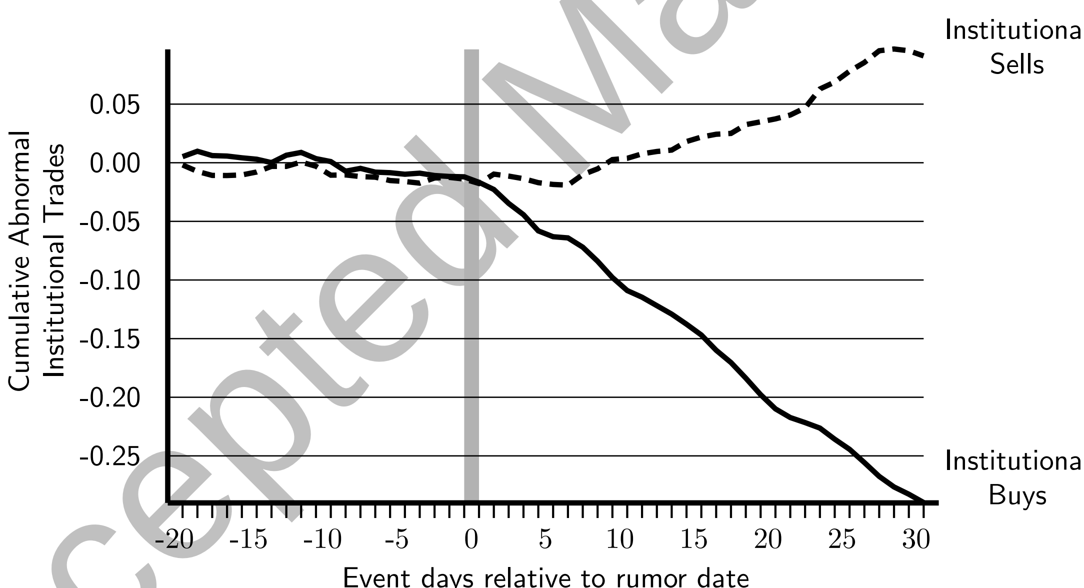


Figure 3
Abnormal shorting utilization by rumor accuracy in event time
This figure presents the time series relative to the date of the first publication of a merger rumor of average abnormal shorting utilization of targets in rumors. Shorting utilization is the fraction of shares on loan divided by shares lendable for short positions. Abnormal short utilization is daily short utilization minus the average daily utilization over the six month period that ends twenty days before the rumor is published. Data are from Markit from 2006 to 2011.



(A) Institutional Buy-Sell Imbalance



(B) Institutional Fraction of Total Volume

Figure 4
Daily institutional trading in event time
This figure presents the time series, relative to the date of the first publication of a merger rumor, of the cumulative abnormal institutional buy-sell imbalance of trading in targets of rumors in panel A and institutional buys and sells normalized by total volume in panel B. Buy-sell imbalance is the daily number of shares purchased minus the number of shares sold by all institutional traders in the ANcerno database, normalized by the sum of buys and sells. Abnormal buy-sell imbalance is the daily buy-sell imbalance minus the average daily imbalance over the six-month period that ends 20 days before the rumor is published. Panel B presents the cumulative abnormal fraction of total trades by institutions in ANcerno. The fraction of total volume is all ANcerno trades/CRSP trading volume, where ANcerno trades are separated into buys and sells. Abnormal trading volume is normalized as above. Data are from the period 2000–2011.

Table 1
Rumor articles by calendar period

	All articles	Scoop articles	Percentage of scoops	Percentage of SDC bids	Rumors that came true	Percentage that came true
Panel A: Yearly						
2000	155	35	7.0	9.6	12	34.3
2001	123	40	8.0	6.6	19	47.5
2002	185	60	12.0	6.1	15	25.0
2003	130	37	7.4	6.6	9	24.3
2004	58	14	2.8	7.3	4	28.6
2005	184	37	7.4	8.6	18	48.6
2006	185	56	11.2	9.7	15	26.8
2007	279	70	14.0	11.2	25	35.7
2008	214	31	6.2	9.1	9	29.0
2009	131	25	5.0	7.4	7	28.0
2010	393	75	15.0	8.8	26	34.7
2011	105	21	4.2	9.0	8	38.1
Total	2,142	501			167	33.3
Panel B: Monthly						
January	209	46	9.2	7.5	12	26.1
February	134	32	6.4	7.4	12	37.5
March	183	51	10.2	8.9	17	33.3
April	223	47	9.4	8.0	17	36.2
May	246	48	9.6	8.5	20	41.7
June	157	43	8.6	8.9	16	37.2
July	150	36	7.2	8.8	7	19.4
August	100	28	5.6	7.9	4	14.3
September	256	53	10.6	8.0	15	28.3
October	156	42	8.4	8.3	10	23.8
November	185	40	8.0	8.2	17	42.5
December	143	35	7.0	9.6	20	57.1

This table presents counts of rumor articles by year and month. "All articles" includes all rumor articles in the sample. The scoop article is the first article that reports a rumor. "Percentage of scoops" is the fraction of total scoop articles in the sample that were published in a given year or month. "Percentage of bids in SDC" is the fraction of bids in the SDC database where the original public announcement is in a given calendar period. "Rumors that came true" is the number of rumors with a publicly announced bid in the SDC database within one calendar year of the original scoop article. "Percentage that came true" is the number of rumors that came true divided by the number of scoop articles for a given time period.

Table 2
Summary statistics of predictive variables

	Mean	Std. Dev.	Percentile 25th	Percentile 50th	Percentile 75th
Panel A: Target newsworthiness					
Public target	0.884	0.320	1	1	1
Log(Target book assets)	9.394	1.990	8.056	9.472	10.563
Valuable brand	0.156	0.363	0	0	0
Advertising/Assets	0.008	0.024	0	0	0
Industry sales to households	0.384	0.279	0.102	0.371	0.591
Tobin's q	1.731	1.471	0.961	1.199	2.007
R&D/Assets	0.015	0.043	0	0	0
Log(1+Distance)	5.960	2.849	5.101	7.111	8.149
Foreign target	0.251	0.434	0	0	1
Panel B: Journalist characteristics					
Log(Journalist age)	3.612	0.191	3.481	3.584	3.724
Undergraduate major					
Business & Economics	0.098	0.297	0	0	0
Journalism	0.328	0.470	0	0	1
English	0.311	0.463	0	0	1
Political Science	0.192	0.395	0	0	0
History	0.257	0.438	0	0	1
Other	0.098	0.297	0	0	0
College SAT Percentile	83.729	12.438	78	87	95
Expert in target industry	0.554	0.498	0	1	1
New York–based	0.495	0.501	0	0	1
Award winner	0.176	0.381	0	0	0
Gender	0.454	0.498	0	0	1
Panel C: Article characteristics					
Weak modal words (%)	0.745	0.542	0.395	0.677	0.975
Anonymous source	0.920	0.272	1	1	1
Target Comment					
Declined to comment	0.459	0.499	0	0	1
Has conversations	0.010	0.099	0	0	0
Confirmed rumor	0.032	0.176	0	0	0
Denied rumor	0.038	0.191	0	0	0
Couldn't be reached	0.084	0.277	0	0	0
Wasn't asked	0.377	0.485	0	0	1
Merger Stage					

continued on next page

Table 2 - *Continued*

	Mean	Std. Dev.	Percentile 25th	50th	75th
Speculation	0.507	0.500	0	1	1
Preliminary talks	0.086	0.280	0	0	0
In talks	0.269	0.444	0	0	1
Made offer	0.050	0.218	0	0	0
Preparing bid	0.036	0.186	0	0	0
For sale	0.032	0.176	0	0	0
Evaluating bids	0.020	0.140	0	0	0
Rumor in headline	0.848	0.359	1	1	1
Number of bidders mentioned	1.500	1.517	1	1	2
Price mentioned	0.386	0.492	0	0	1
Articles on scoop date (#)	1.729	1.338	1	1	2
Panel D: Newspaper characteristics					
Family-run media company	0.735	0.442	0	1	1
Log(Newspaper age)	4.493	0.966	4.718	4.779	4.927
Log(Newspaper circulation)	13.720	1.071	12.902	14.277	14.554

This table presents summary statistics for the main variables used throughout the paper for 501 merger rumors over the period 2000–2011. Observations are at the scoop article level. The scoop article is the newspaper article that first reports the merger rumor. Journalist characteristics are aggregated from individual journalist characteristics to the scoop article level. Variable definitions are presented in the Appendix.

Table 3
Accuracy rates of journalists and media sources

Journalist/Media source	All articles	Percentage of all articles	Scoop articles	Percentage of scoops	Accuracy rate
Panel A: Journalists					
Dennis K. Berman (WSJ)	71	3.2	24	4.0	62.5
Andrew Ross Sorkin (NYT)	67	3.0	19	3.2	42.1
Nikhil Deogun (WSJ)	27	1.2	13	2.2	53.8
Robert Frank (WSJ)	20	0.9	13	2.2	23.1
Robin Sidel (WSJ)	23	1.0	9	1.5	55.6
Anupreeta Das (WSJ)	21	0.9	7	1.2	42.9
Michael J. de la Merced (NYT)	32	1.4	6	1.0	16.7
Jeffrey McCracken (Bloomberg)	19	0.9	6	1.0	50.0
Anita Raghavan (NYT)	19	0.9	6	1.0	16.7
Suzanne Kapner (NYT)	18	0.8	6	1.0	33.3
Sarah Ellison (DJNS)	16	0.7	6	1.0	33.3
Erica Copulsky (NY Post)	14	0.6	6	1.0	33.3
786 others	1,867	84.3	482	79.9	36.3
Total	2,214	100.0	603	100.0	37.6
Panel B: Media sources					
Wall Street Journal	448	20.9	158	31.5	38.6
Dow Jones News Service	625	29.2	67	13.4	38.8
New York Times	219	10.2	38	7.6	28.9
Reuters News	73	3.4	26	5.2	19.2
New York Post	95	4.4	24	4.8	37.5
Barron's	38	1.8	15	3.0	26.7
NYT Blogs	59	2.8	12	2.4	16.7
Bloomberg	16	0.7	10	2.0	80.0
Boston Globe	38	1.8	8	1.6	25.0
Financial Times	15	0.7	8	1.6	62.5
Los Angeles Times	7	0.3	6	1.2	16.7
94 others	509	23.8	129	25.7	25.2
Total	2,142	100.0	501	100.0	33.3

This table presents publishing activity for journalists and media sources in the sample from 2000 to 2011. The scoop article is the newspaper article that first reports the merger rumor. Accuracy rate is the fraction of scoop articles in which a formal takeover bid is made for the target firm within one year. The total number of articles and scoops for journalists exceeds the total number for media sources because some articles include multiple authors. Abbreviations in parentheses indicate the most recent newspaper that employed a journalist. *WSJ* indicates *Wall Street Journal*, *NYT* indicates *New York Times*, *DJNS* indicates *Dow Jones News Service*, and *NY Post* indicates *New York Post*.

Table 4
Target characteristics in rumors versus actual mergers

	Rumors	All mergers	Large public targets	U.S. merger targets
Public target (%)	88.42	38.10***	100.00***	36.87***
		(< 0.001)	(< 0.001)	(< 0.001)
Log(Target book assets)	9.39	7.45***	9.39	6.99***
		(< 0.001)	(0.998)	(< 0.001)
Valuable brand (%)	15.57	0.48***	2.27***	0.44***
		(< 0.001)	(< 0.001)	(< 0.001)
Advertising/Assets (%)	0.83	0.14***	0.10***	0.38***
		(< 0.001)	(< 0.001)	(< 0.001)
Industry sales to households (%)	38.41	30.53***	34.42***	29.83***
		(< 0.001)	(0.003)	(< 0.001)
Tobin's q	1.73	1.64	1.24***	1.85
		(0.299)	(< 0.001)	(0.189)
R&D/Assets (%)	1.55	0.61***	0.43***	1.26
		(< 0.001)	(< 0.001)	(0.152)
Foreign (%)	25.15	66.08***	76.99***	0.00***
		(< 0.001)	(< 0.001)	(< 0.001)

This table presents average characteristics of target firms in the rumor sample compared with targets in actual mergers. Targets in actual mergers are taken from SDC over the period 2000 to 2011 and exclude mergers that are in the rumor sample. The column denoted "All mergers" includes private, public, and subsidiary mergers of targets across the globe, where deals must be worth at least $250 million. Mergers in the column denoted "Large public targets" include the subset of public targets where the minimum target book assets is set such that the average firm in the subsample has the same book assets as the average firm in the rumor sample. The column "U.S. merger targets" only includes targets in the United States, but does not constrain size or public status of the target. The numbers in parentheses are p-values from t-tests of the average of each merger column with the rumor column average. Significance at the 1%, 5%, and 10% levels is indicated by ***, **, and *.

Table 5
Target abnormal event returns and reversals

	All	Rumor came true: Yes	Rumor came true: No	Difference
Panel A: Rumor publication date				
Day 0	4.271***	6.865***	3.029***	3.835***
	(< 0.001)	(< 0.001)	(< 0.001)	(0.001)
Panel B: Run-up period				
Days [−20, −1]	3.093***	5.267***	2.045**	3.222
	(< 0.001)	(0.003)	(0.041)	(0.112)
Days [−10, −1]	2.483***	3.850***	1.824**	2.026
	(< 0.001)	(0.005)	(0.029)	(0.205)
Days [−5, −1]	2.116***	2.903***	1.736***	1.167
	(< 0.001)	(0.005)	(0.003)	(0.320)
Panel C: Post-rumor period				
Days [+1, +5]	−0.850	1.313*	−1.885**	3.198***
	(0.177)	(0.090)	(0.027)	(0.005)
Days [+1, +10]	−1.395*	1.422	−2.743***	4.165***
	(0.062)	(0.107)	(0.007)	(0.002)
Days [+1, +20]	−0.849	2.480	−2.442**	4.922**
	(0.371)	(0.114)	(0.039)	(0.012)
Panel D: Complete period				
Days [−20, +20]	6.550***	14.736***	2.633	12.103***
	(< 0.001)	(< 0.001)	(0.112)	(< 0.001)

This table reports average cumulative abnormal returns in percentages. Abnormal returns are raw returns minus the CRSP value-weighted index. Rumors that came true are those in which an official takeover announcement was made within one year of the first report of the rumor in the press. The numbers in parentheses are p-values. Significance at the 1%, 5%, and 10% levels is indicated by ***, **, and *.

Table 6
Rumor accuracy and stock returns: Target newsworthiness

Dependent variable:	Rumor comes true			Return day 0
	(1)	(2)	(3)	(4)
Day 0 return		4.458***		
		(< 0.001)		
Estimated deal likelihood			0.282***	
			(< 0.001)	
Estimated announcement return				0.083
				(0.393)
Log(Target book assets)	−0.287***	−0.264***	−0.279***	
	(0.004)	(0.005)	(0.007)	
Valuable brand	−0.500*	−0.382*	−0.461*	−0.042***
	(0.078)	(0.082)	(0.077)	(0.006)
Advertising/Assets (%)	−0.068**	−0.073***	−0.069**	0.001
	(0.035)	(0.007)	(0.021)	(0.796)
Industry sales to households	0.323	0.377	0.432	0.002
	(0.315)	(0.351)	(0.196)	(0.901)
Tobin's q	−0.045	−0.022	−0.030	−0.005**
	(0.206)	(0.574)	(0.420)	(0.020)
R&D/Assets (%)	0.007	−0.005	0.004	0.003**
	(0.847)	(0.860)	(0.910)	(0.022)
Log(1+Distance)	0.006	−0.013	−0.010	0.003**
	(0.729)	(0.475)	(0.619)	(0.020)
$\text{Returns}_{(-5,-1)}$	0.012	1.500	0.553	−0.279**
	(0.990)	(0.271)	(0.548)	(0.034)
Industry fixed effects	yes	yes	yes	no
Year fixed effects	yes	yes	yes	no
Observations	399	399	399	399
Pseudo/Adjusted R^2	0.103	0.138	0.126	0.112

continued on next page

Table 6 - *Continued*

This table examines the relationship between target newsworthiness and (1) the likelihood that a rumor comes true and (2) the target stock returns on the day the rumor is first published. Columns 1–3 present fixed effect logit regression coefficients in which the dependent variable is a dummy variable equal to one if a merger rumor came true within one year of the first rumor date. Column 4 presents OLS regression coefficients where the dependent variable is the rumor target's abnormal stock return on the date the first rumor article is published. Abnormal returns are calculated as the target returns minus the CRSP value-weighted index return. Variables are defined in the appendix. Industry fixed effects include Fama-French 17 industry codes. Standard errors are clustered at the industry level and p-values are reported in parentheses. Significance at the 1%, 5%, and 10% levels is indicated by ***, **, and *.

Table 7
Rumor accuracy and stock returns: Journalist characteristics

Dependent variable:	Rumor comes true			Return day 0
	(1)	(2)	(3)	(4)
Day 0 return		8.569***		
		(< 0.001)		
Estimated deal likelihood			0.496***	
			(0.006)	
Estimated announcement return				0.103*
				(0.085)
Log(Journalist age)	1.378**	1.247**	1.232**	0.038
	(0.027)	(0.015)	(0.039)	(0.370)
Undergraduate degree				
Business & Economics	0.223	0.104	0.102	0.025
	(0.697)	(0.842)	(0.853)	(0.211)
Journalism	1.177***	1.214***	1.265***	0.007
	(0.003)	(0.002)	(< 0.001)	(0.638)
English	0.069	−0.051	0.074	0.019
	(0.851)	(0.901)	(0.844)	(0.338)
Political Science	0.272	0.105	0.281	0.019
	(0.506)	(0.798)	(0.488)	(0.128)
History	0.626	0.450	0.547	0.050
	(0.153)	(0.399)	(0.298)	(0.135)
Other	0.548	0.334	0.557	0.021
	(0.280)	(0.442)	(0.217)	(0.408)
College SAT percentile	0.016	0.013	0.013	0.000
	(0.261)	(0.292)	(0.352)	(0.655)
Expert in target industry	0.592*	0.627**	0.572*	0.003
	(0.073)	(0.047)	(0.086)	(0.847)
New York–based	0.440*	0.082	0.295	0.037***
	(0.052)	(0.785)	(0.333)	(0.004)
Award winner	0.209	0.591*	0.308	−0.057*
	(0.469)	(0.094)	(0.361)	(0.069)
Gender	−0.512	−0.425	−0.495	−0.012

continued on next page

Table 7 - *Continued*

Dependent variable:	Rumor comes true			Return day 0
	(1)	(2)	(3)	(4)
	(0.143)	(0.193)	(0.158)	(0.231)
Returns$_{(-5,-1)}$	2.532***	5.007***	3.360***	−0.243***
	(0.001)	(< 0.001)	(< 0.001)	(0.001)
Log(Target book assets)	−0.340***	−0.290***	−0.334***	
	(< 0.001)	(< 0.001)	(< 0.001)	
Industry fixed effects	yes	yes	yes	no
Year fixed effects	yes	yes	yes	no
Observations	296	296	296	296
Pseudo/Adjusted R^2	0.201	0.264	0.233	0.121

This table examines the relationship between journalist characteristics and (i) the likelihood that a rumor comes true and (ii) the target stock returns on the day the rumor is first published. Columns 1–3 present fixed effect logit regression coefficients in which the dependent variable is a dummy variable equal to one if a merger rumor came true within one year of the first rumor date. Column 4 presents OLS regression coefficients where the dependent variable is the rumor target's abnormal stock return on the date the first rumor article is published. Abnormal returns are calculated as the target returns minus the CRSP value-weighted index return. Variables are defined in the appendix. Industry fixed effects include Fama-French 17 industry codes. Standard errors are clustered at the industry level and p-values are reported in parentheses. Significance at the 1%, 5%, and 10% levels is indicated by ***, **, and *.

Table 8
Rumor accuracy and stock returns: Article characteristics

Dependent variable:	Rumor comes true			Return day 0
	(1)	(2)	(3)	(4)
Day 0 return		4.582***		
		(< 0.001)		
Estimated deal likelihood			0.273***	
			(0.001)	
Estimated announcement return				0.099
				(0.118)
Weak modal words (%)	−0.834***	−0.877***	−0.857***	0.005
	(0.001)	(0.003)	(0.001)	(0.444)
Anonymous source	0.367	0.306	0.366	0.019***
	(0.603)	(0.676)	(0.602)	(0.004)
Target response				
Has conversations	−0.193	−0.040	−0.041	−0.007
	(0.718)	(0.942)	(0.940)	(0.529)
Confirmed rumor	1.234***	0.975*	1.119**	0.076**
	(0.010)	(0.081)	(0.026)	(0.016)
Denied rumor	−0.999	−1.068	−0.984	0.007
	(0.271)	(0.246)	(0.250)	(0.655)
Couldn't be reached	0.446	0.259	0.403	0.048***
	(0.209)	(0.458)	(0.250)	(< 0.001)
Wasn't asked	−0.101	−0.158	−0.055	0.014
	(0.828)	(0.731)	(0.903)	(0.165)
Merger stage				
Preliminary talks	0.728**	0.599*	0.620*	0.014
	(0.019)	(0.061)	(0.060)	(0.312)
In talks	1.319***	1.340***	1.319***	−0.003
	(0.001)	(< 0.001)	(0.001)	(0.785)
Made offer	0.611	0.604	0.684	−0.003
	(0.505)	(0.487)	(0.509)	(0.820)
Preparing bid	0.648	0.582	0.536	0.017
	(0.488)	(0.520)	(0.571)	(0.613)

continued on next page

Table 8 - *Continued*

Dependent variable:	Rumor comes true			Return day 0
	(1)	(2)	(3)	(4)
For sale	0.361	0.459	0.590	−0.015
	(0.461)	(0.385)	(0.201)	(0.406)
Evaluating bids	0.965	1.057	0.959	−0.022
	(0.455)	(0.407)	(0.449)	(0.299)
Articles on scoop date (#)	0.164**	0.057	0.107	0.022**
	(0.043)	(0.437)	(0.176)	(0.031)
Rumor in headline	0.002	0.019	−0.051	0.000
	(0.998)	(0.978)	(0.940)	(0.992)
Number of bidders mentioned	0.098**	0.101**	0.101***	−0.001
	(0.011)	(0.036)	(0.010)	(0.763)
Price mentioned	0.667***	0.674***	0.634***	0.008
	(0.002)	(0.005)	(0.005)	(0.487)
Returns$_{(-5,-1)}$	0.013	1.229*	0.442	−0.266*
	(0.982)	(0.068)	(0.496)	(0.056)
Log(Target book assets)	−0.273***	−0.242**	−0.260**	
	(0.004)	(0.020)	(0.015)	
Industry fixed effects	yes	yes	yes	no
Year fixed effects	yes	yes	yes	no
Observations	372	372	372	372
Pseudo/Adjusted R^2	0.224	0.244	0.237	0.202

This table examines the relationship between article characteristics and and (1) the likelihood that a rumor comes true and (2) the target stock returns on the day the rumor is first published. Columns 1–3 present fixed effect logit regression coefficients in which the dependent variable is a dummy variable equal to one if a merger rumor came true within one year of the first rumor date. Column 4 presents OLS regression coefficients where the dependent variable is the rumor target's abnormal stock return on the date the first rumor article is published. Abnormal returns are calculated as the target returns minus the CRSP value-weighted index return. Variables are defined in the appendix. Industry fixed effects include Fama-French 17 industry codes. Standard errors are clustered at the industry level and p-values are reported in parentheses. Significance at the 1%, 5%, and 10% levels is indicated by ***, **, and *.

Table 9
Long-short portfolio returns

	All	Illiquidity		Idiosyncratic risk	
Accuracy model		Low	High	Low	High
Newsworthiness	0.964*	0.181	1.429**	0.190	1.102
	(0.058)	(0.799)	(0.033)	(0.717)	(0.231)
Journalist characteristics	1.014*	0.044	0.585	0.782	0.340
	(0.083)	(0.954)	(0.440)	(0.145)	(0.712)
Article characteristics	0.932*	−0.292	1.004	0.427	1.068
	(0.098)	(0.694)	(0.149)	(0.407)	(0.231)
Journalist fixed effects	0.707	−0.147	1.206*	−0.165	1.009
	(0.200)	(0.846)	(0.065)	(0.738)	(0.239)
Newspaper fixed effects	0.703	−0.012	1.438**	0.221	1.609*
	(0.196)	(0.987)	(0.024)	(0.663)	(0.073)
Journalist and newspaper fixed effects	0.618	−0.314	1.160*	−0.174	1.619*
	(0.295)	(0.711)	(0.080)	(0.728)	(0.067)

This table reports monthly returns from an equally–weighted long-short portfolio, with a long position in targets of rumors that are predicted to be more accurate than the median accuracy rate and a short position in targets with a predicted accuracy less than the median accuracy rate. Accuracy is predicted using the logit models based on the characteristics listed under the column named Accuracy Model, after controlling for the Day 0 return. For each set of characteristics, the predictive model is based on column 2 of the respective table: Table 6 (target newsworthiness), Table 7 (journalist characteristics), Table 8 (article characteristics), Internet Appendix Table 10 (journalist fixed effects), Internet Appendix Table 13 (newspaper fixed effects), and unreported tests that combine journalist and newspaper fixed effects in one model. Firms are added to the portfolio on the day after the rumor is published, so the Day 0 return is not included in the long-short portfolio return. Firms are held for up to one year and the portfolio is rebalanced monthly. If less than five firms meet the criteria for the long or short portfolio, the long-short portfolio return is assumed to be zero. The monthly returns presented here are the daily return compounded 20 days. Newey and West (1987) p-values are in parentheses. We follow Newey and West (1994) to compute lags. Significance at the 1%, 5%, and 10% levels is indicated by ***, **, and *.

Table 10
Run-up and total premium

Dependent variable:	Target CAR (-42,-1)		Target CAR (0,+5)		Target CAR (-42,+5)	
Rumor	0.077***	0.076***	−0.078***	−0.079***	0.000	−0.008
	(< 0.001)	(< 0.001)	(<.001)	(< 0.001)	(0.994)	(0.503)
Completed		0.029**		0.040***		0.072***
		(0.050)		(< 0.001)		(< 0.001)
Majority cash		−0.001		−0.018		−0.017
		(0.806)		(0.248)		(0.253)
Tender offer		0.038***		0.096***		0.141***
		(< 0.001)		(<.001)		(< 0.001)
Leveraged buyout		−0.017**		0.015		−0.015
		(0.046)		(0.213)		(0.110)
Cross-border		0.014		0.002		0.012
		(0.126)		(0.892)		(0.495)
Target takeover defenses		−0.021		0.020		0.003
		(0.212)		(0.144)		(0.907)
Firm-level controls	no	yes	no	yes	no	yes
Industry fixed effects	yes	yes	yes	yes	yes	yes
Year fixed effects	yes	yes	yes	yes	yes	yes
Observations	2431	2431	2431	2431	2431	2431
Adjusted R^2	0.035	0.053	0.063	0.110	0.062	0.113

This table presents fixed effects OLS models of the cumulative abnormal stock returns of merger targets in the three windows: run-up (event window $(-42, -1)$ relative to the public announcement date), announcement $(0, +5)$, and the combined period $(-42, +5)$. Observations include mergers for public U.S. targets that were announced in 2000–2011. Variables are defined in the appendix. All regressions include target size (log of market equity). Firm-level controls include advertising/assets, industry sales to households, R&D/assets, Tobin's q, and a dummy for a valuable brand. The numbers in parentheses are p-values from standard errors clustered at the Fama-French 17 industry level. Significance at the 1%, 5%, and 10% levels is indicated by ***, **, and *.