Review of Finance, 2021, 997–1046
doi: 10.1093/rof/rfaa037
Advance Access Publication Date: 9 December 2020

OXFORD

Financial Media, Price Discovery, and Merger Arbitrage*

Matthias M. M. Buehlmaier[1] and Josef Zechner[2]

[1]Faculty of Business and Economics, The University of Hong Kong, Pokfulam Road, Hong Kong and [2]WU Vienna University of Economics and Business

Abstract

Using merger announcements and applying methods from computational linguistics we find strong evidence that stock prices underreact to information in financial media. A one standard deviation increase in the media-implied probability of merger completion increases the subsequent 12-day return of a long-short merger strategy by 1.2 percentage points. Filtering out the 28% of announced deals with the lowest media-implied completion probability increases the annualized alpha from merger arbitrage by 9.3 percentage points. Our results are particularly pronounced when high-yield spreads are large and on days when only few merger deals are announced.

JEL classification: G11, G14, G34

Keywords: Financial media, Merger arbitrage, Hedge funds, Market efficiency, Mergers and acquisitions

Received November 19, 2019; accepted November 19, 2020 by Editor Joel Peress.

1. Introduction

Whether competitive financial markets can efficiently aggregate information is one of the central questions in finance. Since a consensus seems to have been reached both in theory

* We thank Benjamin Golez, Gerard Hoberg, Byoung-Hyoun Hwang, Gunter Löffler, Micah Officer, and Ying Zhang, seminar participants at the City University of Hong Kong, Lingnan University, University of Hong Kong, University of Ulm, and HKSFA/CFA Institute as well as participants in the SFM 2013 and EFA 2014 conferences and the 2016 Textual Analysis Workshop at Ludwig-Maximilian University of Munich for valuable suggestions and discussions. We are responsible for all remaining errors. The work described in this paper was substantially supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China [Project No. HKU 740612B]. Furthermore, this research has been partially supported by the following grant: HKU Seed Funding for Basic Research. An earlier version of this paper was circulated under the title, "Slow-moving Real Information in Merger Arbitrage."

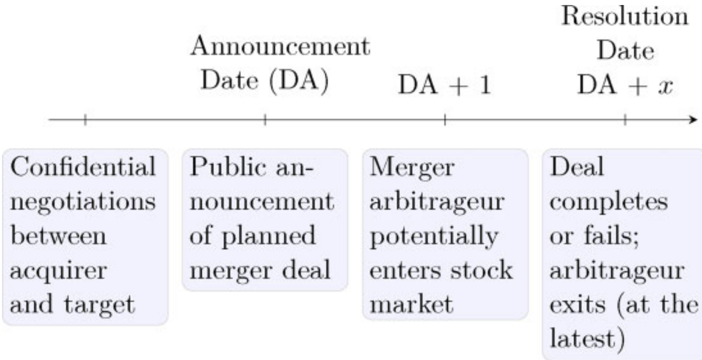

Figure 1. Merger arbitrage timeline.

and in the empirical literature that prices of financial assets are not fully informationally efficient, recent research focuses on understanding the limits of arbitrage and the resulting dynamics of the price discovery. This paper contributes to this literature by applying methods of textual analysis in the context of a well-defined corporate event: corporate mergers. Specifically, we analyze whether texts in financial media provide information about the probability of merger completion which is not already contained in the target stock price.

Analyzing price discovery in a merger context has several advantages. First, in contrast to many other corporate events such as earnings announcements, a "merger" announcement is largely unanticipated by the market and thus the time of information release is well defined. This is so because before the announcement without inside information, it is almost impossible to predict the exact pairing of firms involved and the exact timing of the announcement (see, e.g., Palepu, 1986).[1] Second, the relevant information for the short-term price dynamics of the target is also well defined, namely information about the probability of merger completion. Thus, this information is unambiguously observable *ex post* since there are only two possible outcomes, observable *ex post* without measurement error: either the merger is withdrawn or the two companies complete the merger. Furthermore, this final outcome is known relatively quickly, on average within 3–4 months after the merger announcement. This greatly simplifies our research design compared with other studies that focus on merger synergies, which are difficult to measure and may take years to materialize.

To measure the stock market reaction to new information from the merger announcement, we use a variant of the common merger arbitrage strategy described in Mitchell and Pulvino (2001). Specifically, the merger arbitrage strategy is a risky bet, initiated on the merger announcement date by a deal outsider (the merger arbitrageur), who bets that the merger will complete. In the simplest case, illustrated in Figure 1, the arbitrageur buys the stock of the target company if he believes that the merger is going to complete. If later this belief turns out to be true and the merger indeed completes, the profit is the positive difference between the bid price and the target's stock price (the so-called arbitrage spread). On the other hand, if the merger fails and the companies do not merge, the arbitrageur typically

1 This does not contradict our results in this paper because in contrast to Palepu (1986), we make a prediction *after* the announcement has occurred and we predict a different economic variable, namely whether the merger will complete.

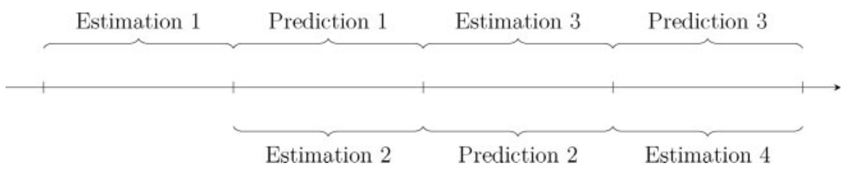

Figure 2. Rolling estimation of textual data. This figure illustrates the timeline of the rolling estimation applied to textual data. Each time period corresponds to one quarter of a year. Each quarter is used for estimating ("fitting") the textual model, and the fitted model is then applied out-of-sample in the following quarter for prediction. For example, the fitted model used for "Prediction 1" in the figure below is estimated in the previous quarter labeled "Estimation 1." The model for "Prediction 2" is estimated using data from "Estimation 2," and so forth.

loses money, since he has to sell the target company's stock at a loss. The distinguishing feature of this strategy is that it is a simple bet on merger completion.

Merger arbitrage is a well-known investment strategy used almost exclusively by institutional investors. Unless there are limits to arbitrage, this market segment should be highly efficient. If we find economically and statistically significant effects of financial media in merger arbitrage, a compelling contribution to our existing understanding of market efficiency and limits to arbitrage can be made.

To estimate the probability of merger completion, we use one of the best-established models from computational linguistics, the naïve Bayes model, and apply it to textual media content from press articles and newswire articles (Lewis, 1998; Manning and Schütze, 1999). The naïve Bayes model is simply based on Bayes' theorem with the added simplifying assumption that all words occur statistically independent from each other. Specifically, we analyze the words in each press article to calculate a media content measure, which relates this textual information to the probability of merger completion. We thus call our media content measure the "media-implied completion probability" and use this term synonymously with "media content" throughout the paper. In symbols, $P(\text{completion}) = f(\text{media content})$, that is, we model the probability of merger completion directly as a function of textual media content, with the function f represented by the naïve Bayes model. We estimate this model on a quarterly rolling basis (see Figure 2) and then apply it out-of-sample to predict mergers in the following quarter. To avoid any look-ahead bias, we estimate the model after discarding ongoing merger attempts whose resolution dates lie in a subsequent quarter (i.e., mergers whose outcomes cannot be known at the time of estimation).

We find strong evidence that stock prices underreact to information in financial media both in event time as well as in calendar time tests. Our cross-sectional regressions show that a one standard deviation increase in the media-implied probability of merger completion results in an increase of 1.2 percentage points in the subsequent 12-day stock return (2.2 percentage points per month) of the target firm. Furthermore, we find that when we vary the holding period after the announcement day from 1 to 12 days, the return effects of a given increment in media-implied merger completion probability increase monotonically, implying that information indeed is slow-moving.

We find that merger arbitrage becomes significantly more profitable if one uses media information to filter out those announced deals with low completion probability. If one uses media information to eliminate those deals with a media-implied completion probability of less than or equal to 85%, which is equivalent to filtering out approximately 28% of all announced deals, then this increases the annualized risk-adjusted return of the trading strategy [i.e., the "alpha" of the Fama and French (1993) three-factor model] by 9.3 percentage points.

We find that price efficiency relative to media information varies with financial market conditions. Following Axelson *et al.* (2013), we use the Merrill Lynch US High Yield Master II Option-Adjusted Spread to proxy for market conditions and find that media-based profits are particularly large and significant when it is hard for institutional investors to lever up, as indicated by a large high-yield spread. In this case, annualized risk-adjusted returns increase by 11.3% when filtering out deals with low *ex ante* media-implied completion probability. By contrast, such profits decrease significantly or vanish completely when high-yield spreads are small.

Finally, we analyze whether our findings are systematically related to the number of merger announcements made on a particular day. This would be the case if, for example, market participants are subject to limited attention. We therefore distinguish between Mondays when usually more mergers are announced, and other days of the week. On Mondays on average 0.7 mergers are announced, whereas this number is less than 0.4 on other days of the week. We find that on Mondays media information does not help to predict merger arbitrage returns, consistent with capacity limits for financial media to process useful information.

We do not claim that most journalists know more than sophisticated investors. In fact, we do not claim that any individual journalist could trade profitably on the information revealed in her press article because each journalist has a signal that is too noisy. The machine learning algorithm we use for textual analysis on the other hand is able to aggregate information from many such noisy signals to extract economically meaningful knowledge. While straightforward in principle, this is a non-trivial exercise in information collection, analysis, and aggregation. In contrast, an individual journalist with her own fragmented information set cannot easily put into context which part of her textual content is transporting an important message about merger completion.

The media measure for merger completion discussed so far is based on textual media "content," but we also examine an alternative media measure based on media "coverage." Coverage, which counts the number of press articles being released, is fundamentally different to content because it ignores the textual information contained in each press article. As such, it may be easier to manipulate, and prior studies have shown that media coverage may be biased by firms seeking to manipulate their stock price (Ohl *et al.*, 1995; Ahern and Sosyura, 2014), for example, by hiring public relations firms and running media campaigns, or through advertising (Bushee and Miller, 2012; Gurun and Butler, 2012). In this sense, coverage complements our content measure. Coverage may "capture" manipulation attempts, while our content measure is designed to "filter out" manipulation in order to extract fundamental information about merger completion.

We find weaker results for media coverage, consistent with the notion that coverage may be easier to manipulate. For example, our time series regressions show that a one standard deviation increase in media coverage yields an increase in annualized merger arbitrage returns of only 9.6 percentage points, whereas the same change in lagged media content yields a larger increase in annualized returns of 11.3 percentage points. Furthermore, a trading strategy based on media coverage is statistically insignificant, while the same trading strategy based on media content significantly increases annualized alphas by 9.3 percentage points.

These results show that in order to extract slow-moving real information about merger completion from financial media, it is not sufficient to simply use media coverage. Instead, a deeper investigation is required and it is important to examine the content of those press articles using textual analysis.

We also find weak evidence in favor of a certification role of the media. We restrict the construction of our media measures to the top newspapers and top newswires to separately investigate their information content. We find that often these top news sources contribute more novel information to the market, consistent with their certification role to stock market investors.

To the best of our knowledge, this is the first paper to model the probability of an economic outcome as a "direct" function of textual media content. We do not use an indirect dictionary approach, where words are classified according to positive or negative psychological associations or according to some domain-specific categories (e.g., Ahern and Sosyura, 2015). This is particularly important, since it has been shown by Loughran and McDonald (2011) that financial texts may be easily misclassified. We tackle this problem at its root by directly relating each word in a press article to merger completion, which is the central focus of this study. In this sense, the naïve Bayes model allows us to infer, directly from the data, the "meaning" of each word regarding "merger completion." Thus, in contrast to prior studies, we do not impose meaning in a potentially subjective or *ad hoc* way by using dictionaries that according to our results only pick up part of the relevant information about merger completion. Instead, we let the data speak by estimating meaning directly from the data. A consistency check of the most important words and combinations of words (so-called bigrams) identified by our algorithm confirms that these are not random words from data dredging and instead capture economically meaningful textual content about merger completion. Furthermore, we demonstrate that naïve Bayes has a performance comparable to an alternative model that is at the forefront of computational linguistics, referred to in the literature as "random forests."

The remainder of the paper is organized as follows. Section 2 reviews the related literature. We then introduce the merger arbitrage investment strategy in Section 3 and explain how we quantify media information in Section 4. In the following sections, we describe and analyze the data: Section 5 details our data and shows summary statistics, Section 6 presents our regression results, and in Section 7, we investigate the determinants of price efficiency. We then present robustness checks in Section 8 using different methods of textual analysis, including a dictionary approach and the random forests model. Finally, Section 9 concludes.

2. Related Literature

Our paper is related to several strands of literature. First, it is connected to work on the profitability of merger arbitrage strategies. Examples for this literature are Larcker and Lys (1987), (Dukes, Frohlich, and Ma, 1992; Karolyi and Shannon, 1999; Jindra and Walkling, 2004). Although the profitability of merger arbitrage has been decreasing over time (Jetley and Ji, 2010), these papers document that substantial risk premia can be realized from merger-related investment strategies, even when based solely on publicly available information.

Potential explanations for these findings include market inefficiencies such as limits to arbitrage (Baker and Savaşoglu, 2002; Officer, 2007) as well as trading costs and premia for providing liquidity during market downturns (Mitchell and Pulvino, 2001; Mitchell, Pedersen, and Pulvino, 2007). Another explanation proposed by Cornelli and Li (2002) is that merger arbitrageurs have an informational advantage relative to target shareholders because arbitrageurs, hiding among noise traders, know that they bought shares. The good performance of merger arbitrage has also been attributed to hedge funds' better management of downside risk (Cao *et al.*, 2014), to investors underreacting to the passage of time after the merger announcement (Giglio and Shue, 2014), or to probability weighting (Wang, 2017).

Other studies suggest that even prior to the announcement day, some inside information may leak into options markets (Jayaraman, Mandelker, and Shastri, 1991; Cao, Chen, and Griffin, 2005; Chan, Ge, and Lin, 2013; Augustin, Brenner, and Subrahmanyam, 2019). In contrast to this literature, our study does not focus on the profitability of merger arbitrage strategies *per se*, but instead analyzes how fast post-announcement information from financial media gets incorporated in stock prices and whether the results differ across time and/or cross-sectionally.

Second, our paper is also related to the literature on slow-moving information in financial markets. Several papers find that stock prices are sluggish to incorporate new information. One of the most salient findings is described in Huberman and Regev (2001), where a reprint of a more than 5 months old story causes stock prices to soar. While this case might have been an isolated incident, it has been shown using larger samples that indeed some information diffuses slowly into asset prices, with contributions including Klibanoff, Lamont, and Wizman (1998), Chan (2003), Mitchell, Pedersen, and Pulvino (2007), Duffie (2010), Tetlock (2010), Tetlock (2011), Mitchell and Pulvino (2012), Da, Gurun, and Warachka (2014), Manela (2014), Peress (2014), and Scherbina and Schlusche (2015). For example, Peress (2014) documents how information transmission to financial markets is impacted during newspaper strikes. In a broader context, financial media are receiving increasing attention in the literature, including studies that investigate long-term returns to media content strategies, for example. In addition to the papers cited above, recent contributions include Veldkamp (2006); Tetlock (2007); Dyck, Volchkova, and Zingales (2008); Tetlock, Saar-Tsechansky, and Macskassy (2008); Bhattacharya *et al.* (2009); Fang and Peress (2009); Dyck, Morse, and Zingales (2010); Da, Engelberg, and Gao (2011); Engelberg and Parsons (2011); Griffin, Hirschey, and Kelly (2011); Dougal *et al.* (2012); Engelberg, Reed, and Ringgenberg (2012); Gurun and Butler (2012); Kuhnen and Niessen (2012); García (2013); Fang, Peress, and Zheng (2014); Hillert, Jacobs, and Müller (2014); Liu, Sherman, and Zhang (2014); and Solomon, Soltes, and Sosyura (2014). For example, García (2013) shows that stock return predictability using sentiment as a measure of media content is concentrated in recessions. We also study potentially slow-moving information. Compared with most existing studies, mergers provide a more structured setting since they define a clear-cut timeline of information flow, namely the time window between the announcement date and the time of ultimate resolution of the merger, while simultaneously allowing us to study the speed of information diffusion.

Third, our paper is related to the literature on the role of media in mergers and acquisitions more generally. Liu and McConnell (2013) show that managers, who have reputational capital at risk, can be swayed by media attention to abandon value-reducing acquisition attempts. Another channel through which the media can influence merger outcomes is through costly signaling, with the media transmitting fundamental information on whether the merger creates or destroys value (Buehlmaier, 2015). On the other hand, instead of providing novel information, the media can also introduce noise to financial markets, with sensationalist merger rumors published prior to the merger announcement about firms that interest the newspapers' readers (Ahern and Sosyura, 2015). There may also be attempts to manipulate the media in order to sway investors (Ahern and Sosyura, 2014). This is consistent with the notion that public relations and marketing campaigns are effective not only in non-financial contexts, but also when people make "financial" decisions (Ohl *et al.*, 1995; Cook, Kieschnick, and Van Ness, 2006; Bushee and Miller, 2012; Solomon, 2012; Lou, 2014). This study contributes to this literature by identifying a specific information transmission channel through textual media content about merger completion and by examining the information diffusion in the context of mergers. Unlike some other studies, we focus on tradable information on or after the merger announcement and we demonstrate that there is fundamental information being released since stock prices do not revert to their previous levels.

Finally, our paper is related to finance applications of textual analysis outside the area of financial media. For example, a growing literature analyzes the textual content of corporate filings such as 10-K statements (Hoberg and Phillips, 2010; Li, 2010; Loughran and McDonald, 2011; Jegadeesh and Wu, 2013; Hoberg, Phillips, and Prabhala, 2014; Loughran and McDonald, 2014a; Buehlmaier and Whited, 2018), or SB-2/S-1 filings (Hanley and Hoberg, 2010; Loughran and McDonald, 2013). Other sources of textual information that have been analyzed include stock message boards on the Internet, assessing the impact of investor opinions on management announcements, stock market volatility, and trading volume (Antweiler and Frank, 2004; Das and Chen, 2007), central bank announcements (Romer and Romer, 2004; Rosa and Verga, 2007; Lucca and Trebbi, 2011; Rosa, 2011; Loughran and McDonald, 2014b; Schmeling and Wagner, 2019), or rating reports (Löffler, Norden, and Rieber, 2019). While our paper uses similar algorithms to capture textual information as some of these studies, it differs by focusing on the information content of financial media during merger events and on the speed at which this information gets incorporated in stock prices.

3. Merger Arbitrage

This section sets the stage for our analysis by briefly defining merger arbitrage investment strategies and clarifying the terminology surrounding it. As mergers are sometimes announced after trading hours, the merger arbitrage returns are always computed starting from the opening price of the first trading "after" the announcement day. The holding period may vary, but all positions are closed at the latest on the resolution day of the merger, that is, the day of completion or withdrawal. Figure 1 provides an overview of the timeline.

There are two types of merger deals: cash deals and stock deals. The former means the acquirer pays by cash, while the latter means the acquirer pays at least partially with its own shares. In a cash deal, the merger arbitrage strategy consists of buying the target company's stock, so the return is r_{Tar}. In a stock deal, on the other hand, the merger arbitrage strategy may consist of going long the target, or additionally shorting the acquirer's stock, yielding a return of $r_{\mathrm{Tar}} - \delta r_{\mathrm{Acq}}$, where δ is the deal's exchange ratio.

More detailed background information about merger arbitrage mechanics is provided in the Online Appendix available at www.buehlmaier.net.

4. Quantifying Information in Financial Media

This section describes how we capture information about merger completion from financial media. We first identify press articles and newswire articles which mention the names of both the acquirer firm and the target firm in the first 100 words of the article and which have been published in a period that starts with the week prior to the announcement date and ends with the resolution date. While the universe of articles thus covers the whole merger period, our analysis will avoid any look-ahead bias. For example, in cross-sectional regressions where we use returns consisting of the first few trading days after the announcement, we only use media information released up to the announcement day, but not thereafter.

Our main media measure captures the probability of merger completion, calculated from media content. We thereby analyze "all" the words in a press or newswire article, thus analyzing media content. We then relate media content directly to the probability of merger completion, since this is the central determinant of the ultimate profitability of the merger arbitrage strategy. In symbols, we model the *ex ante* probability of merger completion as a function f of media content as follows:

$$P(\text{merger completion}) = f(w_1, w_2, \ldots, w_n), \tag{1}$$

where $(w_1, w_2, \ldots, w_n)$ is the word count vector. Thus, w_i, $i = 1, \ldots, n$, is the count of word i in the given media article. For example, if $w_i = 8$ then word i appears eight times in the press article. We use the $n = 2{,}191$ words that are not stop words and appear in at least 1% of all press and newswire articles to ensure we have sufficient coverage of all important words. (A list of stop words and a description of the procedure to remove sparse words can be found in the Online Appendix in Section A.) Thus, we do not restrict the words to a potentially subjective set of keywords and instead keep all words in order to "let the data speak" and to allow for a more precise estimation of our statistical model in Equation (1).

Following the literature on computational linguistics, the word count vector $(w_1, w_2, \ldots, w_n)$ is a simplified representation of a given press article, disregarding word order or grammar and keeping the count of each word. While some information such as word order is lost, it is a standard way of representing text documents in computational linguistics, because it retains important information about media content (Bird, Klein, and Loper, 2009).

To relate the word count vector $(w_1, w_2, \ldots, w_n)$ to the probability of merger completion, we employ one of the best-established models used in computational linguistics for text classification. In particular, we use the naïve Bayes model to represent the function f in Equation (1). Naïve Bayes is a simplified version of Bayes' theorem, with the simplification consisting of the assumption that words occur statistically independent from each other. Lewis (1998) and Manning and Schütze (1999) provide details and background information about this model. While there are other, more sophisticated models for textual analysis (some of which we analyze for comparison in Section 8), we use naïve Bayes because, after decades of active research, it has become a benchmark that has been shown to deliver good performance across a wide variety of applications (Rish, 2001).

As the name suggests, the naïve Bayes model is simply based on Bayes' theorem to obtain conditional probabilities. It is called "naïve" because of its assumption of independence among words, thus leading to a simplified version of Bayes' theorem. Nonetheless, despite the simplifying independence assumption, the naïve Bayes model performs rather well in practice (Zhang, 2005). Furthermore, it has been shown to perform well even on small training sets as, shown by Brain and Webb (1999) via a variance decomposition and in Beleites *et al.* (2013) via cross-validation.

In some cases, the estimated probabilities from naïve Bayes have been shown to diverge from the true probabilities (Manning, Raghavan, and Schütze, 2008). However, in this paper, we are interested in distinguishing between low- and high-probability mergers, and not in making accurate predictions about the exact success probability of a single merger. To this end, it is sufficient if the actual merger completion probability is an increasing function of the model-implied probability, which we verify in Section 5.3.

To avoid any look-ahead bias, we employ a rolling estimation throughout the paper, using press articles from the previous quarter to estimate the model (on a subset of 520 articles on average each previous quarter), before applying it predictively to all press articles in the current quarter. For an illustration, see Figure 2. If a merger was announced in the previous quarter but the merger resolution (i.e., completion or withdrawal) takes place in a later quarter, we exclude it from the rolling estimation (of the previous quarter) to avoid any look-ahead bias. This methodology is robust to using different time window lengths.

We also evaluate the out-of-sample performance of the naïve Bayes model by running a standard cross-validation on each training set (i.e., the previous quarter from the rolling estimation). Using five-fold cross-validation, we find that there are on average 95% correctly classified instances.[2] We are therefore confident that we identify relevant information about merger completion not only in-sample, but also out-of-sample, and that the precision of the estimates is relatively high.

While Equation (1) analyzes media content for a single media article, we aggregate this media content measure for a given deal-day when "several" press articles are released on that day. To this end, we average the merger completion probabilities of all press articles that pertain to a given merger deal on a given day. Specifically, since we aggregate press articles for a single merger deal (and not for several deals, which could have different firm sizes), we "equal"-weight the completion probabilities. Since each probability is between one and zero, the averaged aggregate media content measure is also between one and zero. Its interpretation thus remains unchanged, so the aggregate media content measure captures the probability of merger completion for a given merger deal on a given day.

In addition to media content, we also employ an alternative media measure, media coverage, that captures the arrival intensity at which new information is provided by financial media. Media coverage is a much simpler measure than content, since coverage ignores the words in each article and instead counts how many press articles are released on a given day about a given merger. For example, if coverage is equal to eight on a given day, it means that eight press articles were published about the merger on that day. To capture the surprises in coverage, we additionally consider an adjusted version of media coverage by subtracting its exponentially weighted moving average (EWMA).

While being simpler than media content, media coverage may also be easier to manipulate (Ahern and Sosyura, 2014). Furthermore, media coverage may be less clearly related to the probability of merger completion than media content, as it will be mostly driven by the intensity at which new information becomes available to financial media.

2 When interpreting this number, one needs to keep in mind that about 90% of all mergers succeed. So, even a simple "rule" that always predicts completion would be right about 90% of the time.

Financial media certainly is not the only potential source of information during mergers. For example, definitive merger agreements have been shown to affect acquisition terminations and renegotiations (Denis and Macias, 2013). However, merger agreements can be filed with a substantial time lag relative to information in financial media, which renders them less useful for our research question. For these reasons, we focus on financial media to extract relevant merger-related information in a timely manner.

5. Data, Summary Statistics, and First Results

Section 5.1 describes our data sources, Section 5.2 provides summary statistics, and Section 5.3 shows preliminary results.

5.1 Data

We obtain merger-related information from Thomson Reuters SDC Platinum. Since we want to analyze whether price relevant information about mergers is contained in financial media and how fast this information is impounded in stock prices, it is crucial to obtain accurate announcement dates. While problems with SDC announcement dates have been reported, especially for years before 1984, it is well-known that SDC is accurate for the time period of our sample (Barnes, Harp,]and Oler, 2014).

We include the SDC categories "disclosed value mergers and acquisitions," "tender offers," and "exchange offers," while we exclude "undisclosed value mergers and acquisitions" since the value of the merger deal should be known. We then apply the following screens to the merger data based on screening variables provided by SDC. The merger's announcement date must be in the 11 years between January 1, 1999, and December 31, 2009, which ensures that we include both the dot-com bubble and the financial crisis of 2007–08. This choice of the sample period also ensures that we are not excluding pending merger deals with unknown status because all deals announced in that time period have either been completed or were withdrawn at the time we downloaded the data. This is important because it would otherwise induce a sample-selection bias and make the merger arbitrage strategy of this paper not implementable. Both the target and the acquirer have to be public companies to ensure that stock market data are available. We remove all non-US companies to avoid cross-border mergers, since national interests instead of economic forces often dominate the discussion in those deals (Dinc and Erel, 2013). We also exclude challenged deals, where a second bidder makes an offer to buy the target after the merger announcement has been made. Challenged deals make up less than 2% of the sample and are thus unlikely to make a significant empirical difference, even though the inclusion of challenged deals would make the textual analysis more powerful because (i) there are typically more press articles released and (ii) the articles typically have richer information content. Nonetheless, we exclude challenged deals on conceptual grounds because the competitive nature of auctions can lead to fundamentally different media dynamics that are not the focus of this study.[3] Following common practice in the literature, we also exclude

3 This is not to say that unchallenged deals have no competition; auctions do take place *before* the announcement of a deal, that is, before a firm enters our sample, but there is no data available on these auctions because they are conducted in strict confidentiality by the participating investment banks.

industries that are subject to strong regulation where market forces often have less influence on outcomes, that is, the industries classified by SDC as "energy and power," "financials," and "government and agencies." Deals where the acquirer CUSIP and target CUSIP are identical are removed to avoid contaminating the sample with self-tenders or recapitalizations. We only keep deals where the deal status is completed or withdrawn, where the acquirer owns more than 50% of the target shares after a merger, and where the acquirer purchases at least 20% of the outstanding shares. Of the merger deals that survive these screens, we keep the largest 1,200 deals, as measured by deal value (i.e., the value of the target).

After having created our sample containing the mergers from SDC, we add data from the financial press from Dow Jones Factiva, stock market data from the Center for Research in Security Prices (CRSP), and accounting data from Compustat. A key challenge is to match our merger data from SDC with data from Factiva. To ensure the highest possible data quality, we manually construct text strings that can be used as search terms in Factiva. We do this separately for the target and acquirer, leaving us with 2,400 manually constructed search terms. For example, we search for "IBM or International Business Machines" to capture several variations by which this company is referred to in the press. We then recombine those search terms by requiring that both the search term for the target and the acquirer occur in the first 100 words of a press article. Downloading articles that appear later than the 7 days preceding a merger's announcement date and before a merger's resolution date leaves us with a sample of 130,589 press articles from Factiva. After merging all databases we are left with 1,107 mergers.

To be able to separately investigate the information content of newspapers containing the largest number of articles, we define a group called "top newspapers" that contains the union of the top four newspapers and the top four domestic newspapers, as measured by the number of merger-related articles. This set is given by The Wall Street Journal (2,873 articles), The New York Times (1,621 articles), Financial Times (1,613 articles), The Globe and Mail (956 articles), The Washington Post (757 articles), and eWeek (743 articles). For comparison, we also consider newswires. Since the top newswire, Dow Jones News Service (15,298 articles), already contributes more articles than the top newspapers combined, we only focus on the Dow Jones News Service when analyzing top newswires.

5.2 Summary Statistics

The most important variable for the return of merger arbitrage strategies, deal status, is shown in Table I. Almost 90% of all mergers complete. However, there is still a 10.1% chance that the merger gets withdrawn, in which case a merger arbitrage strategy typically incurs substantial losses. It is therefore of great economic importance to have an *ex ante* information of the likelihood of merger completion.

Table II shows that the unconditional media content measure, which is supposed to measure the probability of merger completion based on media information, closely matches the unconditional probability of merger completion from the previous Table I. The former's mean is 87.0%, while the latter is 89.9%, which shows that the media's *ex ante* estimate of merger completion is "on average" consistent with the fact that *ex post* 89.9% of mergers

Table I. Summary statistics of deal characteristics

The variable "Deal status" indicates whether the acquirer merged with the target. "Stock seal" indicates whether the acquirer paid for the merger using its own stock or cash only. If stock deal is "yes," at least 50% of the merger consideration offered is in the form of acquirer equity. The variable "Unsolicited" denotes whether the acquirer made an offer without prior negotiation with the target.

Variable	Levels	Observations	%
Deal status	Withdrawn	112	10.1
	Completed	995	89.9
Stock deal	No	529	47.8
	Yes	578	52.2
Unsolicited	No	1,037	93.7
	Yes	70	6.3

complete. This does not imply, however, that media correctly predict "individual" merger outcomes, which is a separate question investigated later.

The remaining contents of the table are as expected, with targets holding more cash and having a higher book to market ratio than acquirers (Jensen, 1986), a positive stock market reaction on the announcement date for the target and a negative reaction for the acquirer, an average premium of 33% paid by the acquirer over the target's stock price, and the median duration of each merger of approximately 3 months.

Table III investigates media content and media coverage in event-time of a typical merger, that is, during and after the merger announcement. The table shows a clear spike in media activity on the announcement day, with nineteen articles per day released. After the announcement, however, media activity levels off significantly and only four articles per day are released. Interestingly, the top newswire releases more articles per day than the top newspapers on the announcement day, while the top newspapers release more articles per day in the post-announcement period. This is consistent with the premise that firms manage their media coverage further on the announcement day (Ohl *et al.*, 1995). The media-implied completion probability is 87% on the announcement date, in agreement with the earlier fact from Table I that 90% of all mergers complete. This result is partially driven by the top newswire, which releases the largest number of articles and has a media-implied completion probability of 89% on the announcement day, while the top newspapers come in slightly lower at 86%. The standard deviation is 0.21 on the announcement day, while in the post-announcement period this number increases by 57% to 0.33. Together with the fact that fewer articles are released per day in the post-announcement period, this indicates that in total the media-implied completion probability is able to absorb more information from media on the announcement day than afterward. Furthermore, the completion probability after the announcement decreases to 82%. This result is in accordance with the fact that in our sample, successful deals on average get resolved 11 days sooner than withdrawn deals. The post-announcement sample therefore contains on average more deals that will fail, which gets reflected in a lower media-implied completion probability. The Online Appendix provides additional summary statistics in Section G.

Table II. Summary statistics of continuous variables

The variable "Content$_{DA}$" denotes the probability of merger completion, calculated from textual media content on the announcement day. Coverage$_{DA}$ is the media coverage surprise on the announcement day. Section 4 details the construction of these media measures. The number δ denotes the merger's exchange ratio, and r_{Tar}, r_{Acq}, r_{Mkt}, and r_f are the target's return, the acquirer's return, the stock market's return, and the return on the risk-free rate starting with the opening price on the first trading day after the announcement day until 12 days afterward. These returns capture the returns to merger arbitrageurs. Following the literature, the announcement date is excluded from the return calculation since merger arbitrageurs are assumed to open their stock trading positions after the announcement day. The remaining variables are firm characteristics such as cash to total assets, book to market, and size ($millions), followed by announcement day returns, the merger's duration in days, and the premium paid by the acquirer. All returns are log-returns and the premium is also expressed using logarithms.

Variable	Min	P_{25}	Mean	Median	P_{75}	Max	Std. Dev.
Content$_{DA}$	0.00	0.83	0.87	0.99	1.00	1.00	0.21
Coverage$_{DA}$	−0.42	4.86	17.73	9.68	19.08	649.56	34.51
r_{Tar}	−0.97	−0.02	−0.01	0.00	0.02	0.75	0.12
$r_{Tar} - r_f$	−0.97	−0.02	−0.01	0.00	0.02	0.75	0.12
$r_{Tar} - r_{Mkt}$	−0.98	−0.04	−0.01	0.00	0.04	0.69	0.12
$r_{Tar} - \delta r_{Acq}$	−0.97	0.00	0.01	0.00	0.02	0.70	0.09
Tar. cash/total assets	0.00	0.04	0.28	0.20	0.47	0.99	0.26
Acq. cash/total assets	0.00	0.04	0.21	0.13	0.32	0.93	0.22
Tar. B/M	0.00	0.22	0.52	0.42	0.69	5.52	0.45
Acq. B/M	0.01	0.17	0.40	0.31	0.54	3.28	0.34
Tar. size	2,965.04	126,593.50	1,666,167.28	376,215.75	1,125,051.18	106,213,318.06	5,895,458.97
Acq. size	23,700.25	886,775.22	23,736,267.96	3,292,817.12	16,451,593.36	531,153,308.38	54,332,026.58
$r_{DA,Tar}$	−1.02	−0.02	0.01	0.00	0.03	0.55	0.08
$r_{DA,Acq}$	−0.40	−0.03	−0.01	0.00	0.02	0.31	0.06
Deal duration	0.00	65.50	114.81	92.00	138.00	1063.00	82.03
Premium	−1.93	0.17	0.33	0.32	0.49	1.30	0.28

Table III. Summary statistics of media content and media coverage

This table splits up the timeline of a merger into two periods: the announcement date (Panel A) and the post-announcement period (Panel B). The post-announcement period begins on the first trading day after the announcement and ends on the merger's resolution date, when the merger either completes or is withdrawn. The media-implied completion probability (i.e., media content) is calculated from textual analysis of press articles. In all panels, the media-implied completion probability and media coverage are further divided into all news sources, the top newspapers, and the top newswire. The top newspapers consist of The Wall Street Journal, The New York Times, Financial Times, The Globe and Mail, The Washington Post, and eWeek. The top newswire is Dow Jones News Service. More details about the construction of the media measures are available in Section 4.

			Percentile			
	Mean	Std. Dev.	25th	50th	75th	Obs.
Panel A: Media information during announcement date						
Number of articles per day	19.42	31.15	6	11	21	958
Number of top newspaper articles per day	2.29	2.08	1	1	3	159
Number of top newswire articles per day	3.86	5.47	1	2	4	898
Media-implied completion probability	0.873	0.210	0.83	0.99	1.00	958
Media-implied completion probability from top newspapers	0.859	0.310	1.00	1.00	1.00	159
Media-implied completion probability from top newswire	0.890	0.250	0.98	1.00	1.00	898
Panel B: Media information after announcement date						
Number of articles per day	4.17	9.33	1	2	4	25,426
Number of top newspaper articles per day	2.01	2.19	1	1	2	3,807
Number of top newswire articles per day	1.50	1.47	1	1	1	7,211
Media-implied completion probability	0.821	0.330	0.82	1.00	1.00	25,426
Media-implied completion probability from top newspapers	0.757	0.393	0.51	1.00	1.00	3,807
Media-implied completion probability from top newswire	0.812	0.357	0.92	1.00	1.00	7,211

5.3 First Results

This section provides a first set of findings related to the two media information measures.[4] Figure 3 shows time series plots for both media content and media coverage. Recall that the media content measure captures the media-implied probability of merger completion. The plot shows clearly that media content differs depending on whether the merger is going to complete at the end, with *ex post* completed mergers having an *ex ante* media-based completion probability consistently above 75%, with a slightly rising trend over time. On the other hand, merger arbitrageurs would like to avoid deals that will not complete, and media content already picks up this information *ex ante*. For withdrawn deals, the probability of completion continually drops as time goes on, providing a clear signal that the merger is less likely to complete. While this plot does not tell us how quickly this information is revealed relative to the stock market reaction, it does tell us that already on the announcement day, media content provides an informative signal about merger completion because the solid line is above the dashed line on this day. Media content is thus an informative

4 The Online Appendix contains additional results in Section B, such as the most important words picked up by the linguistic model.

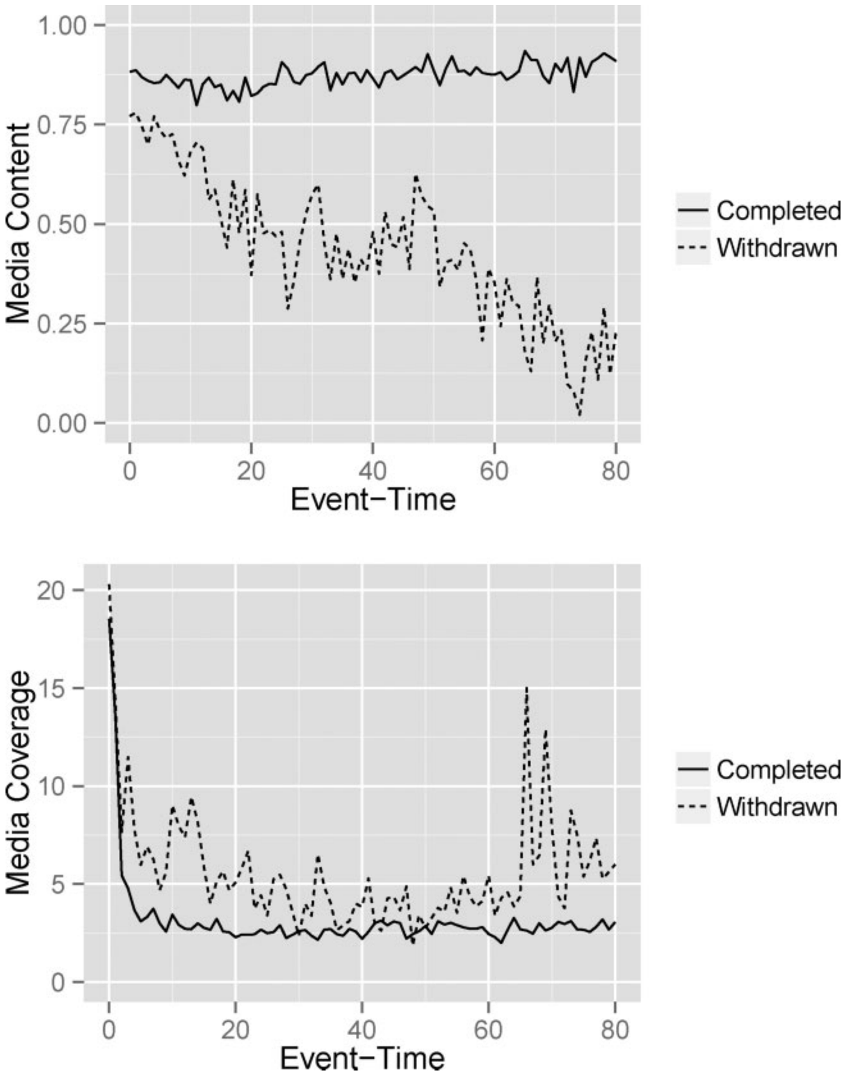

Figure 3. Media measures in event-time. This figure shows the event-time dynamics for the media content and media coverage measures, both split up by whether the merger will complete in the end. Media content is the media-implied probability of merger completion constructed from textual analysis of press articles and newswire articles, while media coverage is the number of press and newswire articles released. Section 3 contains details about the construction of the media measures.

signal on the announcement day about whether the merger will complete. Also, later it picks up information about merger completion, potentially faster than the stock market.

The second plot in this figure shows media coverage, split up by completed and withdrawn merger deals. A clear spike is visible at time zero for both types of deals, showing that the announcement day is particularly important as more press articles and newswire articles get released. After the announcement day, media activity decreases as the surprise of the announcement wanes off. It appears that completed deals receive slightly less media coverage than withdrawn deals, consistent with potential media manipulation attempts of "bad" deals that will later be withdrawn (Ahern and Sosyura, 2014), or alternatively, because withdrawn deals have higher information intensity due to negative news coming out over time, correcting the initially upbeat merger announcement.

Before investigating its relation to stock prices, we provide evidence that the media content measure in fact predicts actual *ex post* merger completion and thus captures the information it is supposed to measure. The purpose of this test is simply to check whether the textual model works as intended. Therefore, at this first stage of the analysis, we are interested only in the predictive power of media without including further control variables (additional controls are included in later tests).

To this end, we use a probit model to regress an *ex post* merger completion dummy on the media content measure, using only press articles released on the announcement day and discarding articles released thereafter. The timing is important because the media measure on the announcement day uses information released on average more than one hundred days "before" the actual merger completion takes place. Furthermore, as in the rest of this paper, to estimate the media measure itself, we use a rolling window (instead of the whole sample consisting of all mergers) to ensure that this is an out-of-sample test and that no look-ahead bias occurs.

Specifically, we use a probit regression of the standard form

$$P(\textit{ex post}\text{ completion}) = \Phi(\beta_1 + \beta_2 \times (\textit{ex ante}\text{ media content})),$$

where P denotes probability, Φ is the standard normal distribution, and β_1 and β_2 are the regression coefficients shown in Table IV. We find that the media content estimate is significantly positive, confirming that media indeed pick up information about actual *ex post* merger completion, even when we restrict media to the announcement date and discard media articles from thereafter.

6. Merger Arbitrage and Information in Financial Media

In this section, we investigate the central question of this study: Do financial media contain fundamental information not yet incorporated in stock prices? As discussed above, our main variable of interest is the probability of merger completion, calculated from textual media content, to which we refer to as "media content."

6.1 Cross-Sectional Merger Arbitrage Results

In Table V, we relate merger arbitrage returns to the probability of merger completion, constructed from textual media content, and media coverage, adjusted for coverage surprises. The cross-sectional regressions in this table are predictive in the sense that all independent variables are known before the dependent variables. In particular, the media variables are constructed using data from the announcement day, while dependent variables are cumulative merger arbitrage returns from the opening price of the first trading day after the announcement day until 12 trading days later (Baker and Savaşoglu, 2002).

While the holding period of 12 trading days is chosen with hindsight, our results stay robust to other holding periods (see also Figure 4). While 12 trading days are "short" compared with the standard merger arbitrage strategy that holds until merger resolution, it is "long" for our purpose, since we only focus on media information released on the announcement day and do not include any press articles released thereafter. If markets are

Table IV. Ex ante media content and *ex post* merger completion

This table shows a probit regression of an *ex post* merger completion dummy on *ex ante* media content. The probit model is of the standard form

$$P(ex\ post\ \text{completion}) = \Phi(\beta_1 + \beta_2 \times (ex\ ante\ \text{media content})),$$

where P denotes probability, Φ is the standard normal distribution, and β_1 and β_2 are the regression coefficients shown in the table below. The media content measure captures the media-implied probability of merger completion, estimated from textual analysis of press and newswire articles. For this regression, we only keep articles released on the merger's announcement date, and discard all later articles. It is therefore a highly predictive regression because merger completion (or withdrawal) on average takes place 3–4 months after the announcement. This regression is a consistency check to verify that the (*ex ante*) media content measure captures information about actual (*ex post*) merger completion. Numbers in brackets show *z*-statistics. Stars indicate significance at 10%, 5%, and 1%.

	Ex post merger completion
Intercept (β_1)	0.41**
	(2.07)
Ex ante media content (β_2)	1.02***
	(4.42)
McFadden R^2	0.03
Nagelkerke R^2	0.04
Number of observation	960

efficient, this information should be "immediately" incorporated into asset prices on the day after the announcement day, so it should not matter how many days we include after the announcement, as long as we include at least the first day. At any rate, we will report results for holding periods shorter and longer than 12 trading days below and also in Section 6.2.

We consider target stock returns, excess target stock returns (over the risk-free interest rate and over the stock market's return), and returns from the long-short investment strategy described earlier in Section 3. We find that the probability of merger completion, constructed from media content, strongly predicts target stock returns over the following 12 days, while media coverage strongly predicts the long-short investment strategy, that is, coverage affects the acquirer's stock reaction. In particular, a one standard deviation increase in the media-implied probability of merger completion yields an increase in the target stock return of 1.2 percentage points over the subsequent 12 days. Assuming 22 trading days per month, this corresponds to an increase in monthly returns by 2.2 percentage points, an economically as well as statistically meaningful number. Similarly, a one standard deviation decrease in media coverage yields a long-short return that is 1.3 percentage points larger over the subsequent 12 days, corresponding to 2.3 percentage points on a monthly basis.

While these are event-date returns and it is not possible to earn them on a continuous basis, there are approximately eight merger announcements per month in our sample. Merger events thus occur frequently enough to make the increase in event-time returns economically relevant.

Table V. Predicting merger arbitrage returns cross-sectionally

This table shows predictive regressions of merger arbitrage returns on the probability of merger completion (calculated from textual media *content*), EWMA-adjusted media *coverage* surprises, and control variables. The media measures are constructed using data from the announcement day, while dependent variables are cumulative stock returns from the opening price of the first trading day *after* the announcement day until 12 trading days later. The regressions are predictive in the sense that independent variables are known before the dependent variables. Numbers in brackets show *z*-values. The symbols *, **, and *** indicate significance at 10%, 5%, and 1%, respectively, based on White standard errors. Coefficients are multiplied by 100 for readability.

	r_{Tar}	$r_{\text{Tar}} - r_{\text{f}}$	$r_{\text{Tar}} - r_{\text{Mkt}}$	$r_{\text{Tar}} - \delta r_{\text{Acq}}$
Intercept	−8.836	−8.847	−6.436	6.562
	(−1.512)	(−1.514)	(−1.257)	(0.938)
$\text{Content}_{\text{DA}}$	5.586***	5.554***	4.494**	−0.894
	(2.857)	(2.838)	(2.492)	(−0.335)
$\text{Coverage}_{\text{DA}}$	0.004	0.005	0.007	−0.036**
	(0.524)	(0.578)	(0.885)	(−2.165)
log(Tar. cash/total assets)	−0.610**	−0.602**	−0.552**	−0.029
	(−2.492)	(−2.460)	(−2.282)	(−0.141)
log(Acq. cash/total assets)	−0.008	−0.002	−0.140	0.249
	(−0.030)	(−0.007)	(−0.527)	(1.079)
log(Tar. B/M)	0.103	0.103	0.179	0.038
	(0.242)	(0.242)	(0.434)	(0.130)
log(Acq. B/M)	0.181	0.210	−0.118	−0.327
	(0.261)	(0.303)	(−0.180)	(−0.569)
log(Tar. size)	−0.058	−0.064	−0.099	−0.250
	(−0.160)	(−0.178)	(−0.301)	(−0.585)
log(Acq. size)	0.217	0.218	0.128	−0.068
	(0.888)	(0.892)	(0.549)	(−0.357)
$r_{\text{DA,Tar}}$	−7.011	−7.075	−4.919	2.248
	(−0.639)	(−0.644)	(−0.479)	(0.240)
$r_{\text{DA,Acq}}$	22.769*	22.903*	20.515*	−5.679
	(1.913)	(1.922)	(1.810)	(−0.595)
Unsolicited = Yes	2.247*	2.279*	1.345	0.191
	(1.813)	(1.837)	(1.196)	(0.163)
Stock deal = Yes	−0.975	−0.979	−1.134*	1.816*
	(−1.391)	(−1.396)	(−1.682)	(1.706)
Premium	1.702	1.694	1.088	−0.882
	(0.800)	(0.796)	(0.554)	(−0.503)
R^2	0.036	0.036	0.033	0.034
Number of observation	910	910	910	624

In Table VI, we repeat the analysis from the previous Table V, with the difference that media data are restricted to be from top news sources. For the top newswire Dow Jones News Service in columns 1 and 2, the coefficients of media content and media coverage increase in absolute value compared with Table V, where all news sources were used. Likewise, the goodness of fit of the regressions increases, as measured by R^2. These results

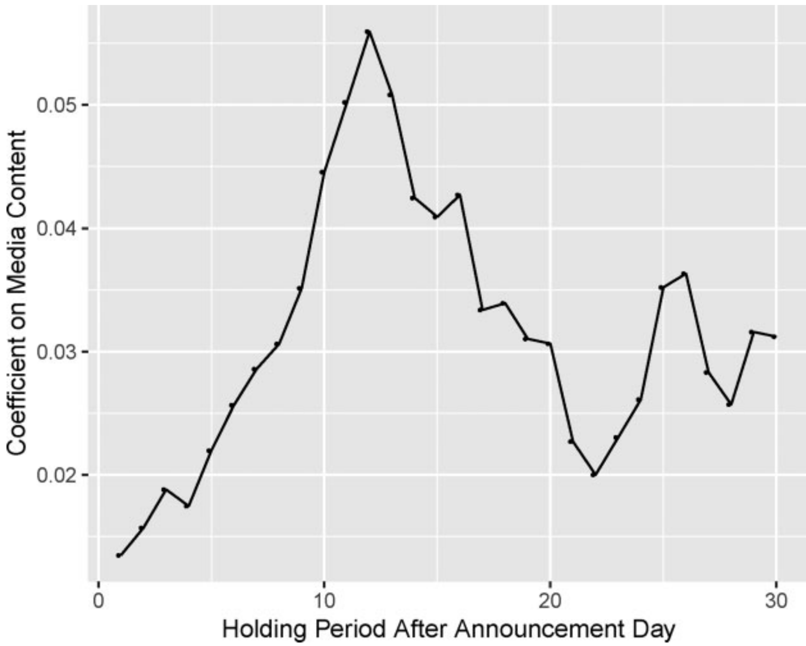

Figure 4. Slow-moving information. This figure shows cross-sectional regression results of the long-only merger arbitrage return on the media content measure and controls. The control variables are the same as in Table V. Specifically, the figure shows the regression coefficients of the media content measure as a function of the holding period of the subsequent return. Each dot on the figure corresponds to a separate regression with a different holding period shown on the *x*-axis. The timing for each regression is as follows. On the merger's announcement day, the media content measure is calculated, and for this calculation we only use information up to and including the announcement day (but no information released afterward to avoid look-ahead bias). On the first trading day *after* the announcement day, we start calculating the cumulative merger arbitrage return until the end of the holding period. If markets are efficient, a longer holding period should not result in a larger media content coefficient, since this media information should already be reflected in prices.

show that the media-implied probability of merger completion is driven by the top newswires. On the other hand, restricting attention to the top newspapers in columns 3 and 4 does not yield significant media coefficients. This might be due to the fact that the top newspapers only publish relatively few press articles, which implies that the number of cross-sectional observations drops to less than 200 in Table VI from over 900 in Table V. Despite the small number of observations, however, the goodness of fit of the newspaper regressions are highest among all cross-sectional merger arbitrage return regressions, with R^2s of up to 11.5%, showing that the top newspapers improve the overall fit of the model considerably.

After establishing that media information has predictive power for merger arbitrage returns, we investigate how quickly media information is absorbed in stock prices. To this end, we repeat the baseline regression from Table V with different holding periods for the left-hand side return. While previously in Table V, the holding period was 12 days starting from the first trading day after the merger's announcement day, we now vary the holding period from 1 to 30 days and plot the resulting media content coefficients in Figure 4. If

Table VI. Merger arbitrage returns and top news sources

This table runs the same predictive regressions as in Table V, with the difference that only top news sources are used for the construction of the media measures. The top newswire is Dow Jones News Service ("TNW" in columns 1 and 2), while the top newspapers ("TNP" in columns 3 and 4) consist of The Wall Street Journal, The New York Times, Financial Times, The Globe and Mail, The Washington Post, and eWeek. We only focus on a single newswire for comparison since it alone produces more articles than the top newspapers combined. "LO" are the long-only excess returns $r_{Tar} - r_f$ of the merger arbitrage strategy while "LS" are the long-short returns $r_{Tar} - \delta r_{Acq}$ introduced in Section 3. Numbers in brackets show z-values. The symbols *, **, and *** indicate significance at 10%, 5%, and 1%, respectively, based on White standard errors. Coefficients are multiplied by 100 for readability.

	TNW LO	TNW LS	TNP LO	TNP LS
Intercept	−10.095*	3.678	−4.555	13.989**
	(−1.723)	(0.626)	(−0.403)	(2.184)
Content_{DA}	6.873***	1.191	2.761	0.400
	(3.896)	(0.649)	(1.447)	(0.234)
Coverage_{DA}	0.053	−0.273*	0.087	0.056
	(0.672)	(−1.773)	(0.788)	(0.200)
log(Tar. cash/total assets)	−0.645**	−0.041	−0.820*	−0.245
	(−2.502)	(−0.190)	(−1.660)	(−0.608)
log(Acq. cash/total assets)	0.072	0.307	−0.347	0.118
	(0.253)	(1.267)	(−0.696)	(0.365)
log(Tar. B/M)	0.127	0.101	−0.425	−0.551
	(0.291)	(0.348)	(−0.623)	(−1.107)
log(Acq. B/M)	0.295	−0.354	0.067	−0.020
	(0.409)	(−0.567)	(0.036)	(−0.019)
log(Tar. size)	−0.101	−0.083	−0.467	−1.103***
	(−0.251)	(−0.223)	(−0.707)	(−2.808)
log(Acq. size)	0.258	−0.109	0.470	0.109
	(1.007)	(−0.550)	(0.820)	(0.314)
$r_{DA,Tar}$	−6.556	0.915	−26.674	−6.322
	(−0.622)	(0.093)	(−0.955)	(−0.475)
$r_{DA,Acq}$	19.287	−8.965	72.783**	−0.358
	(1.532)	(−0.877)	(2.174)	(−0.033)
Unsolicited = Yes	3.087**	1.079	0.648	1.092
	(2.238)	(0.868)	(0.248)	(0.512)
Stock Deal = Yes	−0.880	1.925*	−2.344	0.718
	(−1.207)	(1.790)	(−1.455)	(0.572)
Premium	1.507	−1.182	−4.322	−4.335
	(0.680)	(−0.656)	(−0.761)	(−1.345)
R^2	0.040	0.043	0.115	0.086
Number of observation	853	584	195	129

markets are efficient, the holding period beyond the first day should not matter since the media information should already be reflected in prices and Figure 4 should plot horizontally. However, we find that the plot increases steadily until it peaks at a 12-day holding period, confirming that it takes approximately 2.5 trading weeks before stock prices fully react to the media content measure. Importantly, independent of the holding period, the media content measure only uses information from the merger's announcement day and no information afterward, so the media's information content is held constant as we increase the return's holding period. Thus, the higher return is not due to more information, but it is the "same" information that takes longer to get absorbed into prices. In summary, we find that information provided by financial media only gradually gets reflected in stock prices over a period of approximately 12 trading days.

After a holding period of 12 days, the figure decreases, but always stays above the initial value it had on day 1. This behavior demonstrates that while there seems to be some over-reaction, there is long-lasting fundamental information released about the merger and reflected accordingly in the market.

6.2 Time Series Tests of Media-Based Merger Arbitrage

After establishing in the previous section that information in financial media has cross-sectional predictive power for merger arbitrage returns, we investigate next its time series properties in calendar-time from several angles. We begin by examining a simple trading strategy. If the *ex ante* probability of merger completion, calculated from textual media content on the announcement day, correctly predicts subsequent arbitrage returns (starting from the opening price of the first trading day after the announcement), one should invest in mergers with a high probability of completion and stay away from the remaining mergers. One way to do this is to choose a media content threshold and invest in a deal if announcement day media content is above this threshold, while staying away from deals below the threshold. For each given threshold, we can thus identify a subset of mergers to invest in. For the identified mergers, we then form a merger arbitrage portfolio and calculate risk-adjusted returns in calendar time.

While being simple, this investment strategy allows us to test whether risk-adjusted arbitrage returns depend on announcement date media information. In particular, if the media content threshold is larger on the announcement date, we exclude more mergers with low probability of completion, and should therefore have higher arbitrage returns, because arbitrage returns are a bet on merger completion. So, if merger arbitrage returns are a function of media content, we should observe higher risk-adjusted arbitrage returns for higher media content threshold levels.

Specifically, to formally test this hypothesis, we calculate risk-adjusted arbitrage returns in calendar time for a range of media content threshold levels and plot them and their *p*-values in Figure 5. We calculate risk-adjusted arbitrage returns from the time series of long-short arbitrage portfolio returns starting with the opening price of the day after the announcement day until 12 trading days later to investigate how the media information released on the announcement day (while excluding media thereafter) is absorbed in stock prices in the following trading days. If several deals are open in the 12 day range, we average them in calendar-time to form the arbitrage portfolio. We then regress these arbitrage returns on the stock market return and the Fama/French factors to control for risk (Fama and French, 1993). The risk-adjusted arbitrage return, or "Jensen's alpha," is the intercept of this regression and the *p*-value comes from the intercept as well. The terms "risk-adjusted returns" and "alphas" are used interchangeably.

We hold each merger for 12 days motivated by Figure 4, but our results are robust to other holding periods, even though returns would be slightly lower. Notwithstanding, it is

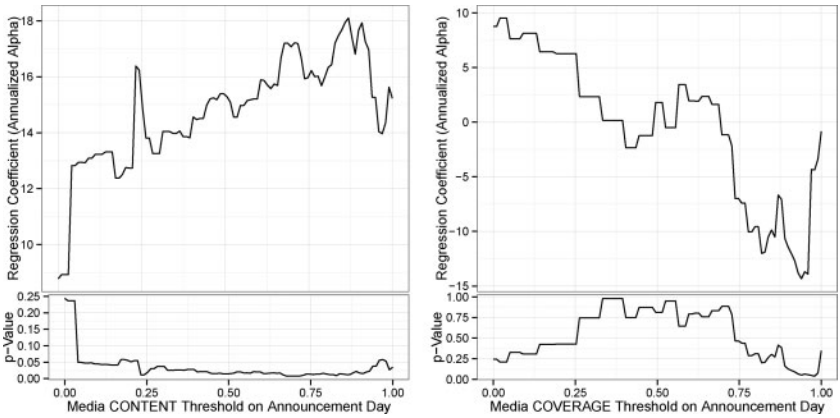

Figure 5. Simple trading strategy. This figure shows risk-adjusted annualized returns ("alphas") from a simple trading strategy that conditions on media content on the announcement day, that is, the media-implied *ex ante* probability of merger completion. If content on the announcement day is above a given threshold level, then starting on the next trading day after the announcement, the merger arbitrageur buys the target and additionally short-sells the acquirer in a stock deal, adjusted by the exchange ratio δ, according to the standard merger arbitrage investment strategy, and holds the deal for 12 trading days to capture potentially slow-moving media information. Otherwise, if the media content measure is below the threshold on the announcement day, he skips investing in this deal. All open deals are averaged in calendar time. The risk-adjusted return ("alpha") for each threshold level is the intercept from regressing the arbitrage portfolio return on the market and the Fama/French factors. The plot on the right-hand side repeats this analysis for media coverage instead of media content.

important to emphasize that we are restricting the profitability of the strategy by not using any media information after the announcement date. If we were to use a richer set of media content including the post-announcement period, we could obtain higher returns using a more sophisticated trading strategy. In this sense, the strategy we consider here can be regarded as a benchmark result contributing a "lower" boundary on possible returns of media-based merger arbitrage strategies.

We find that, consistent with our hypothesis, risk-adjusted arbitrage returns of this trading strategy are an increasing function of the media content threshold, which represents the *ex ante* probability of merger completion based on textual media content. Except for the largest threshold levels, where fewer observations are available, an increase in *ex ante* media content yields a higher arbitrage return, consistent with the notion that merger arbitrage is a bet on merger completion.

The risk-adjusted arbitrage returns are statistically and economically significant. For all threshold levels, the strategy outperforms the benchmark strategy with threshold zero that blindly invests in all mergers, independent of media content. On an annualized basis, the alpha increases from 8.8% when investing in all deals (i.e., the media content threshold is zero) to a maximum of 18.1% when screening out deals that lie below the media content threshold of 0.85. (This threshold is equivalent to filtering out 28% of all announced deals.) This change in alpha corresponds to an increase of 9.3 percentage points in annualized risk-adjusted returns when following a simple trading strategy based on media content.

Our results also show that the alpha is 8.8% when investing in all deals independent of media (i.e., the media threshold is zero). This is consistent with earlier findings (see, e.g., Baker and Savaşoglu, 2002) and a number of possible explanations have been discussed in the literature. We show that these naive merger strategies do not reflect media information and that novel media information takes several days to be absorbed in stock prices, as shown by the increase in calendar-time merger arbitrage returns to 18.1% when trading on financial media information.

Importantly, this increase in returns is accomplished by trading *less*. Because we screen out mergers with a low probability of completion, we obtain a higher return by investing in fewer mergers, which lowers trading costs. In this context, it is important to acknowledge that this trading strategy is very different from a strategy based on factors employed in the asset pricing literature, where one sorts stocks for example once per year on a firm characteristic such as size or value. If we reduce the number of stocks in such a strategy, it is unclear whether trading costs go down, because the remaining stocks might have higher trading costs than the deleted ones. In sharp contrast, the key source of total trading costs in the merger arbitrage strategy is how often one has to trade in and out of deals. If the strategy reduces the number of deals, one has to (i) trade in and out of deals less frequently and (ii) readjust the remaining deals less frequently if several deals are open at any given time. Taken together, our results likely underestimate the impact of media content on merger arbitrage returns since the return spread net of transaction costs increases more than reported here.

We next investigate an analogous strategy for media coverage instead of media content. For media coverage, we find that consistent with earlier cross-sectional results, the risk-adjusted arbitrage returns are a decreasing function of announcement day media coverage. However, the alphas are statistically insignificant, thus contrasting with the stronger results for media coverage.

After investigating the time series properties of a simple trading strategy based on announcement day media content and coverage, we next look into the time series properties of media after the announcement date as well. The timing is now different, since we no longer restrict media to articles published on the announcement day, and instead construct a time series that averages all media content across deals in calendar time from announcement to the merger's resolution. Furthermore, we no longer restrict the portfolio holding period to the first 12 trading days after the announcement and instead include the whole deal duration in the portfolio.

Specifically, in the spirit of Mitchell and Pulvino (2001), we regress time series of arbitrage returns on lagged time series of media measures and controls in Table VII according to the following specification in calendar-time:

$$r_{\mathrm{t}} = \beta_0 + \beta_1 \times \mathrm{content}_{\mathrm{t}-1} + \beta_2 \times \mathrm{coverage}_{\mathrm{t}-1} + \beta_3' \times \mathrm{controls}_{\mathrm{t}}.$$

Here, r_{t} is either the aggregate target excess return or the aggregate long-short merger arbitrage return, media content and coverage are aggregated and lagged by one trading day, and controls include the market and the Fama–French factors. For each calendar day t, aggregation is done across all open deals. Because we are focusing only on the largest deals to begin with, and because the portfolio only holds a small number of stocks (thirty stocks on average for the long-only portfolio and thirty-eight for the long–short portfolio), we aggregate across deals by equal-weighting the arbitrage returns and the media measures on each day.

Table VII. Time series tests of merger arbitrage returns

This table shows time series regressions of merger arbitrage portfolio returns on media measures. The dependent variables are portfolio returns of either the long-only merger arbitrage strategy that buys the target stock after the announcement, or the long-short strategy that in addition to buying the target also short-sells the acquirer in stock deals, adjusted by the exchange ratio δ. The dependent variables include media measures consisting of lagged textual media *content*, capturing the media-implied probability of merger completion, as well as lagged media *coverage*, capturing how many press/newswire articles are released per deal-day. In case more than one merger deal is open at any given date, the returns and the media measures are averaged across all open deals. Coefficients are multiplied by 100 for readability. Numbers in brackets show *t*-statistics. Stars indicate significance at 10%, 5%, and 1%.

	$r_{\mathrm{Tar}} - r_{\mathrm{f}}$	$r_{\mathrm{Tar}} - \delta r_{\mathrm{Acq}}$	$r_{\mathrm{Tar}} - r_{\mathrm{f}}$	$r_{\mathrm{Tar}} - \delta r_{\mathrm{Acq}}$
Intercept	-0.21^{**}	-0.19^{**}	-0.20^{**}	-0.24^{***}
	(−2.20)	(−2.10)	(−2.13)	(−2.59)
$r_{\mathrm{Mkt}} - r_{\mathrm{f}}$	74.17^{***}	24.72^{***}	74.17^{***}	24.69^{***}
	(61.79)	(21.52)	(61.78)	(21.52)
SMB	32.01^{***}	13.89^{***}	32.02^{***}	13.77^{***}
	(12.38)	(5.61)	(12.38)	(5.57)
HML	-22.77^{***}	11.10^{***}	-22.77^{***}	10.95^{***}
	(−9.94)	(5.06)	(−9.93)	(5.00)
Content lagged	0.32^{***}	0.29^{***}	0.32^{***}	0.31^{***}
	(2.82)	(2.69)	(2.81)	(2.86)
Coverage lagged			−0.00	0.01^{**}
			(−0.09)	(2.43)
R^2	0.60	0.15	0.60	0.15
Number of observation	2,820	2,820	2,820	2,820

For all specifications, we find a positive and significant coefficient for the lagged media content measure, capturing the probability of merger completion from textual information. For example, for the long–short merger arbitrage investment strategy, an increase of one standard deviation in lagged media content yields an increase in annualized returns of 11.3 percentage points. As in previous tests in this paper, this is an economically and statistically significant number. The time series variation for media coverage is different compared with earlier tests. While media coverage has a significantly negative sign in the cross-sectional regressions, it now is insignificant for the long-only portfolio and significantly positive for the long–short portfolio, due to the inclusion of the whole duration of the merger. The economic magnitude of this effect is smaller than for lagged media content: a one standard deviation increase in lagged media coverage increases the long–short annualized return by 9.6 percentage points.

Table VIII repeats the regressions from Table VII using media information from only the top news sources. With the exception of top newspapers for the long–short portfolio, all lagged media content variables are significantly positive, thus predicting merger arbitrage returns as expected. For the long-only portfolio, the top newspapers yield an increase in annualized returns of 10.0 percentage points for a one standard deviation increase, while the top newswire yields an increase of 14.4 percentage points. Furthermore, despite fewer

Table VIII. Time series tests of arbitrage returns: top news sources

This table shows the same regressions as in Table VII, with the difference that only top news sources are used in constructing the media measures. "TNP" refers to the top newspapers, "TNW" refers to the top newswire, "LO" refers to the long-only merger arbitrage portfolio that buys the target stock after each announcement, while "LS" refers to the long-short arbitrage portfolio that buys the target and additionally short-sells the acquirer in a stock deal, adjusted by the exchange ratio δ. Coefficients are multiplied by 100 for readability. Numbers in brackets show t-statistics. Stars indicate significance at 10%, 5%, and 1%.

	TNP LO	TNP LS	TNW LO	TNW LS
Intercept	−0.03	−0.01	−0.13***	−0.15**
	(−0.69)	(−0.13)	(−2.59)	(−2.56)
$r_{\mathrm{Mkt}} - r_{\mathrm{f}}$	73.22***	23.13***	73.68***	25.81***
	(50.75)	(14.28)	(67.73)	(20.79)
SMB	35.20***	18.82***	31.70***	12.44***
	(12.06)	(5.74)	(13.60)	(4.67)
HML	−21.41***	16.82***	−23.74***	11.80***
	(−8.36)	(5.85)	(−11.44)	(4.98)
Content lagged	0.11**	0.07	0.21***	0.21***
	(2.20)	(1.18)	(3.85)	(3.38)
Coverage lagged	0.00	0.01	−0.00	0.01
	(0.05)	(0.92)	(-0.10)	(1.21)
R^2	0.64	0.12	0.68	0.16
Number of observation	1,767	1,767	2,469	2,469

observations, the R^2 measures are higher, indicating a better model fit using only a subset of the media data. This shows that information from the top news sources matters more due to better quality or due to a certification effect of the top news sources. In contrast to media content, the media coverage coefficients are all insignificant. This finding again underlines the importance of using standard yet sophisticated textual analysis on press and newswire articles, instead of simply counting the number of articles.

In Table IX, we explore an alternative weighting scheme by using value-weighting in contrast to the equal-weighting previously employed in Table VII. Although our sample only consists of large mergers with high media coverage, it could still be the case that within this group of large firms, our results are driven by the relatively smaller companies. To address this issue, we repeat the analysis of Table VII using value-weighted returns and media measures. The regressions presented in Table IX demonstrate that coefficients of the lagged media content measure, that is, the lagged media-implied completion probability, increase in both magnitude and significance for most specifications compared with equal-weighting in Table VII. A one standard deviation increase in media content (coverage) yields an annualized increase in the long–short returns of 19.1 (13.5) percentage points. This result demonstrates that value-weighting strengthens our narrative, implying that the media-return relation becomes more profitable as size increases. This detail makes sense inasmuch larger firms bear more intense media scrutiny and attention, which makes it easier for our textual analysis to pick up pertinent information about merger completion.

Table IX. Time series tests of value-weighted merger arbitrage returns

This table shows similar regressions as in Table VII, the difference being that the returns and media measures are now value-weighted instead of equal-weighted. Specifically, the dependent variables are the value-weighted merger arbitrage returns, while the independent variables are the market and the Fama–French factors as well as lagged value-weighted media content (the media-implied completion probability) and lagged value-weighted media coverage (the number of press articles released). Returns are value-weighted using the target's market capitalization (Mitchell and Pulvino, 2001), while we use the acquirer's market capitalization for weighting the media measures because media discussion is dominated by the larger acquirer. Coefficients are multiplied by 100 for readability. Numbers in brackets show *t*-statistics. Stars indicate significance at 10%, 5%, and 1%.

	$r_{\text{Tar}} - r_{\text{f}}$	$r_{\text{Tar}} - \delta r_{\text{Acq}}$	$r_{\text{Tar}} - r_{\text{f}}$	$r_{\text{Tar}} - \delta r_{\text{Acq}}$
Intercept	−0.20**	−0.20	−0.21**	−0.25*
	(−2.41)	(−1.55)	(−2.44)	(−1.87)
$r_{\text{Mkt}} - r_{\text{f}}$	81.09***	0.32	81.10***	0.40
	(49.84)	(0.13)	(49.83)	(0.16)
SMB	0.28	2.41	0.25	2.26
	(0.08)	(0.44)	(0.07)	(0.42)
HML	−51.02***	74.58***	−51.04***	74.46***
	(−16.43)	(15.47)	(−16.43)	(15.45)
Content lagged	0.31***	0.31**	0.32***	0.34**
	(3.18)	(2.00)	(3.21)	(2.17)
Coverage lagged			0.00	0.00
			(0.42)	(1.54)
R^2	0.51	0.08	0.51	0.08
Number of observation	2,819	2,819	2,819	2,819

7. Determinants of Price Efficiency

In this section, we explore potential factors that may influence the extent to which financial media contains information not reflected in stock prices and the speed at which this information gets into prices. We first focus in Section 7.1 on financial market conditions that may influence the degree to which arbitrageurs are active. In Section 7.2, we then analyze effects of the number of merger deals announced on a particular day.[5] We then investigate the nature of the information the market is reacting to in Section 7.3 by examining the words most closely associated with merger success and failure.

7.1 Capital Market Conditions, Financial Media, and Stock Prices

In this subsection, we explore whether financial and economy-wide credit conditions may play an important role because they determine the availability of leverage (Axelson *et al.*, 2013). If credit conditions are favorable, it may be easier for merger arbitrageurs to leverage their bets on merger completion and to trade more aggressively on available *ex ante*

5 We also ask if the ownership structure of the target firm is associated with its price efficiency with respect to financial media information, but only find weak evidence, reported in the Online Appendix in Section H.

information about merger completion, including information from financial media. On the other hand, if credit conditions worsen, then trading on information in financial media becomes more difficult and media-based information is therefore arbitraged away more slowly, yielding higher predictive power of media content.

We test this hypothesis by investigating the interaction between lagged media content and a variable capturing debt market condition. Building on Axelson *et al.* (2013), we use the Merrill Lynch US High Yield Master II Option-Adjusted Spread, obtained from the Federal Reserve Bank of St. Louis, to capture economy-wide credit conditions. If this high yield spread is large, credit conditions are unfavorable for arbitrageurs. As a consequence it may be more difficult to take arbitrage positions based on media information, leading to higher predictive power of media content.

To this end, we regress the time series of the merger arbitrage portfolio's returns on an interaction term between the time series of lagged media content and the high yield spread, as well as controls. Consistent with the above hypothesis, Table X shows that the interaction term between the high yield spread and lagged media content is significantly positive. A one standard deviation increase in the high yield spread makes "daily" returns by 0.35 percentage points more responsive to changes in lagged media content, corresponding to 7.6 percentage points per month. This number is economically large, and indicates a strong effect of market conditions on the media-return relationship.

To further investigate how our simple trading strategy from the beginning of Section 6.2 depends on economy-wide credit conditions, we update our previous results by splitting up the trades depending on whether the high yield spread is large or small. In this test, we collect media content on the announcement day (but not thereafter). We then invest in a deal starting on the next day if media content and its captured probability of merger completion is above a given threshold. Deals below that threshold are excluded. We then repeat this strategy for all threshold levels and plot the results separately for both small and large levels of the high yield spread.

Consistent with our hypothesis, Figure 6 shows that the media-based trading strategy works very well when the high yield spread is above its median. For example, in this case, the annualized risk-adjusted return increases by 11.3 percentage points when filtering out deals with an *ex ante* merger completion probability below 90% (which is equivalent to filtering out 39.5% of all announced deals). Thus, when the high yield spread is large, media-based information does not get arbitraged away by investors, possibly due to the unavailability of leverage. If, on the other hand, the spread is below its median, media content on the announcement day loses its predictive power. In these periods, markets seem to react very quickly and stock prices seem to be efficient with regard to information in financial media.

7.2 Merger Monday

Next we investigate whether the media-return relationship is affected by the amount of new information released on a particular day. We see two possible channels through which the amount of information could have an effect. First, there is a literature that documents that investors are subject to limits of attention (see e.g., Kahneman, 1973; Hirshleifer and Teoh, 2003; Barber and Odean, 2008; Da, Engelberg, and Gao, 2011). Thus, if many mergers are announced simultaneously, investors are not able to spend as much attention on each one, thus leading to less efficient prices. If this is true, then we should see a stronger media-

Table X. Availability of leverage to the arbitrageur

This table shows tests investigating how the media–return relationship varies with changes to the availability of leverage to the merger arbitrageur. The dependent variables are the returns of the merger arbitrage portfolios, both the long-only portfolio that buys the target and the long-short portfolio that buys the target and additionally short-sells the acquirer in stock deals, adjusted by the merger's exchange ratio δ. The independent variables include lagged media content, coverage, and controls, as well as an interaction term between lagged content and a standard variable capturing conditions in the credit markets, the Merrill Lynch High Yield Spread. Coefficients are multiplied by 100 for readability. Numbers in brackets show *t*-statistics. Stars indicate significance at 10%, 5%, and 1%.

	$r_{\mathrm{Tar}} - r_{\mathrm{f}}$	$r_{\mathrm{Tar}} - \delta r_{\mathrm{Acq}}$	$r_{\mathrm{Tar}} - r_{\mathrm{f}}$	$r_{\mathrm{Tar}} - \delta r_{\mathrm{Acq}}$
Intercept	-0.27^{**}	-0.30^{***}	0.21	0.29
	(−2.49)	(−2.95)	(1.01)	(1.46)
$r_{\mathrm{Mkt}} - r_{\mathrm{f}}$	74.17^{***}	24.69^{***}	74.14^{***}	24.65^{***}
	(61.78)	(21.52)	(61.82)	(21.53)
SMB	32.01^{***}	13.76^{***}	31.94^{***}	13.67^{***}
	(12.38)	(5.57)	(12.36)	(5.54)
HML	-22.73^{***}	10.99^{***}	-22.77^{***}	10.95^{***}
	(−9.92)	(5.02)	(−9.94)	(5.01)
Content lagged	0.34^{***}	0.33^{***}	−0.24	-0.39^{*}
	(2.99)	(3.07)	(−0.98)	(−1.65)
Coverage lagged	−0.00	0.01^{**}	−0.00	0.01^{**}
	(−0.03)	(2.49)	(−0.17)	(2.31)
High-yield spread	0.01	0.01	-0.06^{**}	-0.08^{***}
	(1.31)	(1.43)	(−2.37)	(−3.13)
Content lagged * high-yield spread			0.08^{***}	0.10^{***}
			(2.68)	(3.48)
R^2	0.60	0.15	0.60	0.16
Number of observation	2,820	2,820	2,820	2,820

return relationship on such days. Alternatively, limits of attention may be affecting the quality of the news media reports on each merger. Here, the premise is that the analysis and evaluation of new corporate information by media journalists takes time and effort. Thus, when many mergers are announced simultaneously, the precision of information reflected in news media may be lower, and therefore the media-return relation becomes weaker or should even temporarily reverse.

To explore these alternative hypotheses, we investigate the effects of Merger Mondays, which are characterized by a larger number of mergers being announced and more news articles released. Figure 7 shows that on Mondays, on average 0.7 mergers are announced, while this number is less than 0.4 on all remaining days of the week, on Fridays even dropping to 0.2. New announcements are also reflected in increased media activity with a distinctive spike of press articles released on Monday. The reasons are that the details of mergers are often finalized over the weekend and announced on Monday morning, or that merger announcements are held on purpose until Monday morning, hoping for a full week of positive press following the announcement. In any case, more mergers accumulate over

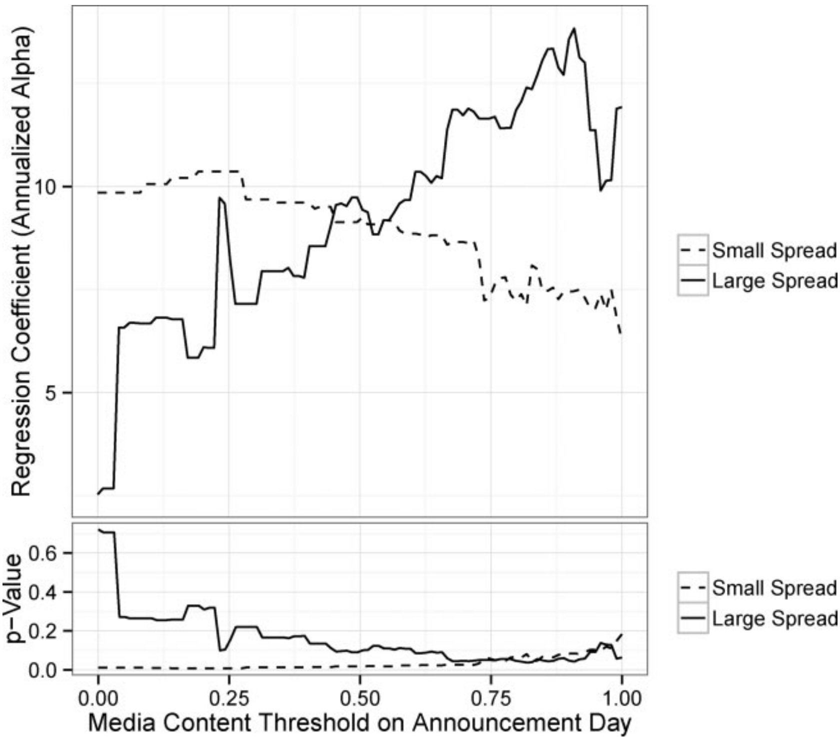

Figure 6. Returns depend on availability of leverage. This figure shows the same plots as in Figure 5, with the difference that the data is split up depending on whether the Merrill Lynch High Yield Spread is above or below its median. Reflecting conditions in credit markets, this spread captures the availability of leverage to the arbitrageur.

the weekend and are released on Mondays, resulting in a burst of new activity due to more information coming out.

While it is well-known that Merger Mondays exist and therefore more mergers are announced, it is important to keep in mind that it is "not" known which mergers are going to be announced, that is, who merges with whom. This means that merger arbitrageurs know in advance that more information will be coming out on Monday, but they do not know in advance which information it is going to be. In this sense, as in the rest of this paper, merger announcements can be treated as unanticipated and exogenous, as it is well known that *ex ante* before the announcement and without inside information it is almost impossible to exactly predict the pairing of firms and the timing of the announcement (see, e.g., Palepu, 1986).[6]

We use the fact that more mergers are announced as a proxy for either less precise media information or fewer active arbitrageurs per deal. Table XI investigates whether the media information released on Mondays has weaker predictive power for merger arbitrage returns. We run two sets of analyses as follows.

6 In contrast to Palepu (1986), we are making predictions *after* the announcement has occurred and we are predicting a different economic variable, namely merger completion.

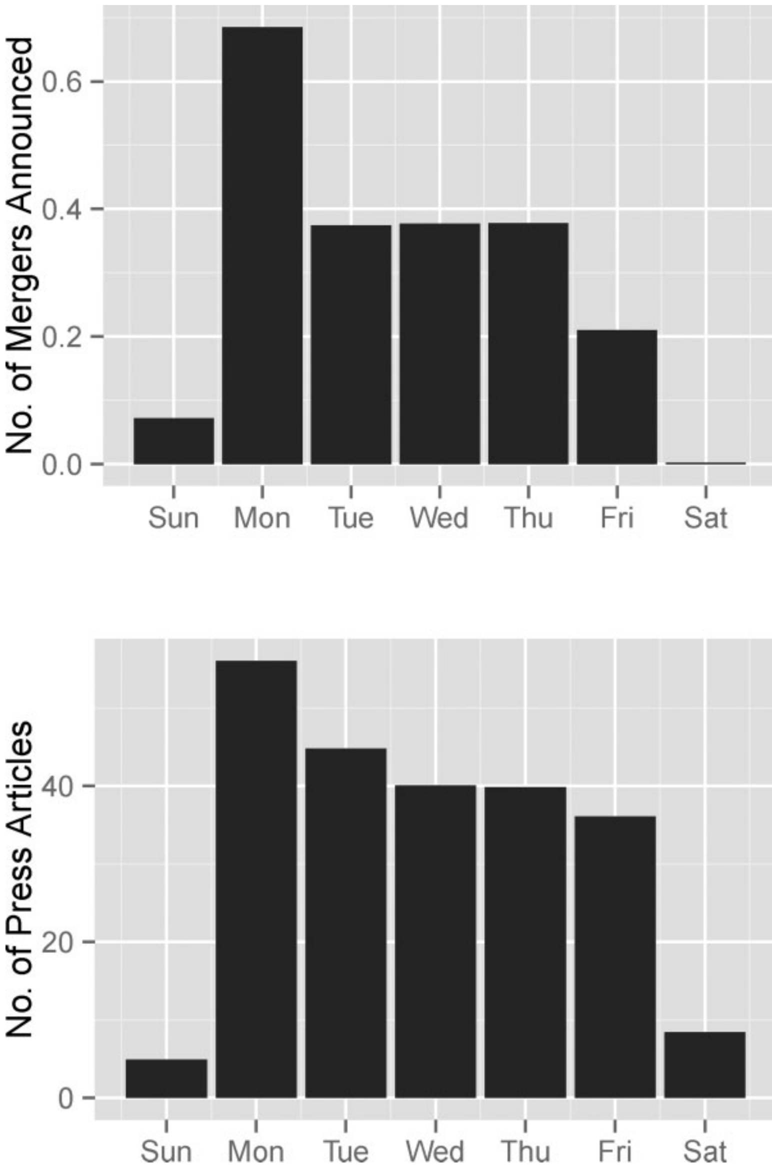

Figure 7. Merger Monday. These plots show for each day of the week the number of announced merger deals and the number of press articles released discussing previously announced merger deals. Merger Mondays can occur for several reasons, one being that details of the merger are finalized over the weekend and announced on Monday, while another reason is holding the M&A announcement on purpose for Monday morning, hoping for a full week of positive press following the announcement.

First, we restrict the sample to the day following Monday (i.e., Tuesday) because we are interested in how "lagged" media variables are related to returns. In other words, we use the subset of data that correspond to media information from Monday and returns from Tuesday. We find that media content coefficients switch sign and become negative,

Table XI. Media during merger Mondays

This table shows the effect of Merger Mondays on the strength of the media–return relationship. Merger Mondays have more mergers announced and more press articles printed (Figure 7), releasing more information in limited time. Since we are interested in "lagged" media content for trading purposes, we focus on the "Tuesday" portfolio returns, because these are the relevant returns if we are interested in the information published by financial media on "Monday," captured as lagged media content. The first two columns therefore show the subsample of merger arbitrage returns on Tuesdays for both the long-only and the long-short strategy (with lagged media content corresponding to Monday). The third and fourth columns show the whole sample with an interaction term added for returns on Tuesdays (i.e., when lagged content is from Monday). Coefficients are multiplied by 100 for readability. Numbers in brackets show *t*-statistics. Stars indicate significance at 10%, 5%, and 1%.

	$r_{\text{Tar}} - r_{\text{f}}$	$r_{\text{Tar}} - \delta r_{\text{Acq}}$	$r_{\text{Tar}} - r_{\text{f}}$	$r_{\text{Tar}} - \delta r_{\text{Acq}}$
Intercept	0.44**	0.25	−0.36***	−0.35***
	(2.15)	(1.13)	(−3.40)	(−3.42)
$r_{\text{Mkt}} - r_{\text{f}}$	73.91***	24.50***	74.09***	24.65***
	(30.26)	(9.42)	(61.81)	(21.51)
SMB	39.66***	10.04*	31.99***	13.66***
	(6.96)	(1.65)	(12.39)	(5.53)
HML	−20.17***	9.10*	−23.04***	10.72***
	(−4.43)	(1.88)	(−10.06)	(4.90)
Content lagged	−0.46*	−0.20	0.51***	0.43***
	(−1.87)	(−0.79)	(4.03)	(3.59)
Coverage lagged	−0.00	0.01	−0.00	0.01*
	(−0.36)	(1.45)	(−0.36)	(1.90)
Tuesday			0.81***	0.60***
			(3.45)	(2.67)
Content lagged * Tuesday			−0.96***	−0.64**
			(−3.43)	(−2.36)
R^2	0.64	0.14	0.60	0.16
Number of observation	577	577	2,820	2,820

consistent with the hypothesis that more and potentially less precise information released makes it more difficult to generate arbitrage profits based on information published in financial media on that day. As a matter of fact, using media information on this day actually hurts performance. Second, we use the whole sample but add an interaction term that captures media information released on Monday. Again, because we use lagged media variables, this dummy is for Tuesday's returns. Consistent with our hypothesis, we find that a one standard deviation increase in media content on a Monday significantly lowers annualized long–short returns by 4.8 percentage points.

Taken together, the evidence in this subsection suggests that financial media information is getting noisier when a large number of different merger deals are reported. Thus, the results are more supportive for the second hypothesis laid out above, namely that limits of attention lead to less accurate information in financial media on such days.

7.3 Nature of Textual Information

To provide more information about the nature of textual information contained in the financial press, we examine the words most closely associated with a merger deal's success and failure. If financial media provide additional signals to market participants about the merger deal, then the words should reflect elements of the deal that are indeed associated with the merger's success or failure. If on the other hand the top words are common words or legalese, it would be more consistent with an irrational response to media "hype" during the merger. Finally, top words related to risk could be consistent with an efficient market but a misspecified asset pricing model not capturing this kind of risk.

To provide richer texture on these questions, we use the fact that the naïve Bayes model is based on Bayes' rule, which allows us to build two lists of words that have the highest conditional probabilities associated with merger success and failure, that is, words with the highest probabilities $P(\text{word}_i|\text{Class} = \text{Success})$ and $P(\text{word}_i|\text{Class} = \text{Fail})$. Words appearing in both lists are excluded in order to isolate the ones most closely associated with success or failure only.

The words displayed below are shown in their root form according to the Porter (1980) stemming algorithm, which is a standard preprocessing step in computational linguistics to increase accuracy. For example, our textual analysis treats "completion" and "complete" in the same way, as both are represented by their root form "complet."

The words most closely associated with merger success are given by: addit, agre, agreement, annual, becom, common, communic, complet, cost, custom, data, debt, develop, give, growth, inform, intern, lead, major, maker, nasdaq, net, pay, purchas, research, state, term, and total.

The words most closely associated with merger failure are given by: advertis, believ, bid, even, giant, googl, issu, just, media, microsoft, much, onlin, possibl, potenti, propos, public, reject, repres, right, rival, sinc, statement, still, street, takeov, unsolicit, work, and yahoo.

The success words are associated with merger details such as whether the parties are in "agreement," whether there are discussions and "communications" about the merger's "completion," "costs," "data" (presumably fundamental data), merger financing ("debt"), future plans to "develop" the merged companies, a hopefully good flow of "information" (e.g., how management keeps shareholders informed), "growth" prospects, whether they are industry "leaders," whether there is support form "major" shareholders, terms of "payment," how they "purchase" the target company, analyst "research" opinions, and whether the "terms" of the merger are discussed in the press.

The failure words, on the other hand, are associated with whether the merger deal has to be "advertised," whether there are uncertainties about the deal ("beliefs," "possible," "potentially," "proposal"), whether the "bid" is discussed (e.g., when the bid is perceived not to be high enough), when deal financing is a concern as securities have to be "issued," whether there is potentially controversial reporting in the "media," when the deal is "rejected" or when there are discussions about whether the offer is "sincere," when there are legal or voting issues (shareholder "rights"), when there is a potential "rival," when the parties are making "statements" (e.g., because there is disagreement about certain parts of the deal), when "the street" believes the deal might not be good (e.g., "the street" as a synonym for Wall Street is often mentioned when there is a negative stock price reaction), when there is an "unsolicited" "takeover," when there are question whether the deal "works," or when certain companies are mentioned such as "Yahoo," "Google," or "Microsoft," that is, the discussion makes reference to companies having a history of frequently making high-profile acquisitions.

Next, to better understand the context of words deemed important by naïve Bayes, we analyze not only single words, but also “bigrams,” which consist of two-word sequences, for example “agree to.” Specifically, we add all combinations of bigrams that occur in our news articles to the previous data set of single words (so-called “unigrams”) and re-estimate the naïve Bayes model on the data set containing both unigrams and bigrams.

The unigrams and bigrams most closely associated with merger success are given by: agre, agreed to, agreement, approv, billion in, by th, complet, custom, develop, earn, expected to, in cash, is expect, million in, monday, network, offic, pay, per shar, presid, purchas, quarter, secur, system, term, the acquisit, the new, to clos, unit, and use.

The unigrams and bigrams most closely associated with merger failure are given by: believ, bid, bid for, come, continu, creat, day, director, for a, friday, get, interest, internet, it i, made, microsoft, move, now, premium, propos, reject, say, take, takeov, the off, to a, unsolicit, well, world, would b, and yahoo.

The first finding is that most terms associated with merger outcomes are unigrams, not bigrams. (Bigrams occur as well, but less often so.) Although bigrams carry potentially more meaning, our results show that in our data set, only few bigrams are more important than unigrams. This is not surprising, as most of the possible two-word combinations do not occur in the data at all. And those that do exist in the data rarely show up consistently in different mergers. However, if a bigram shows up consistently, it has strong power to predict the merger outcome, for example, “per share.”

More importantly, the second finding is that the terms (unigrams and bigrams) make economic sense and thus enhance the prior analysis involving only unigrams. A clear picture emerges that allows us to distinguish between merger deals that are likely to succeed or fail.

Specifically, successful deals are characterized by terms indicating that many of the merger details have been settled by the acquirer and target before the announcement. For example, “by the” (often followed by the word “agreement”), “agreed to,” “expected to” (be completed), “in cash” (referencing the method of payment), “is expected,” “million in” (often followed by a reference to synergies or costs), “per share” (referring to pro forma earnings per share of the merged entity or the purchase price), “the acquisition,” “the new” (often followed by a reference to the merged entity), and “to close” (the deal) are associated with merger success.

On the other hand, failed deals are associated with bigrams expressing uncertainty regarding the planned acquisition. For example, “bid for,” “for a” (often followed by “takeover,” indicating an unsolicited nature of the deal), “the offer” (indicating discussions about the offer price), “to a” (often in the context of “subject ‘to a’ number of terms/conditions”), and “would be” indicate merger failure. These bigrams express uncertainty regarding the planned acquisition, that a “takeover” is “unsolicited,” or that there is still ongoing “bidding for” the target, that is an auction with other bidders or a bidding war.

Altogether these findings are consistent with market inefficiencies resulting from investor inattention and/or limits to arbitrage. The top words are indeed related to what is happening in the deal and convey information consistent with merger success or failure. This result sheds light on the specific mechanism that enables the financial media to transmit fundamental news about the merger to market participants. It is further consistent with Ohl *et al.* (1995), who demonstrate that media often repeat the talking points of deal

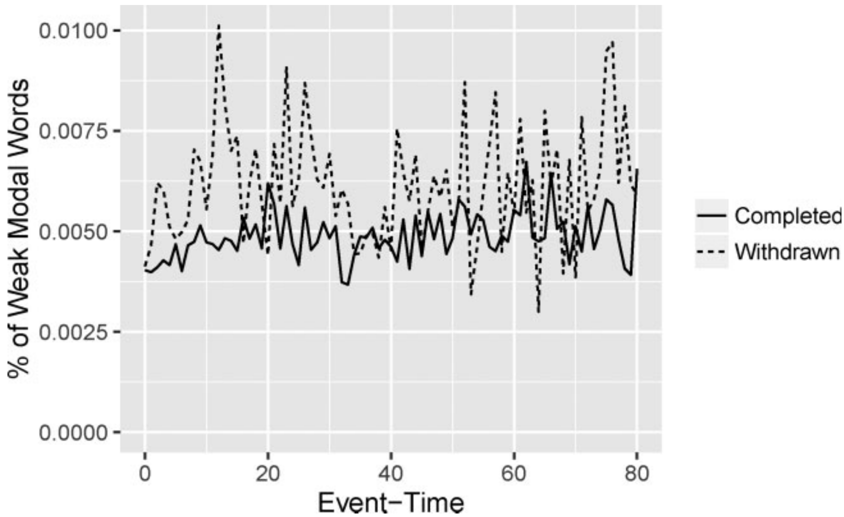

Figure 8. Weak modal words in event time. This figure shows the percentage of weak modal words in event time of the merger. The solid line corresponds to announced mergers that will later complete, while the dotted line is for announced deals that will later be withdrawn.

insiders such as CEOs and that media are thus an important source of new information for investors in a merger environment.

8. Robustness Checks

In this section, we investigate alternative models of analyzing the textual content of newspaper and newswire articles, weak modal words, tonal words, random forests, and logistic regression. We do this to investigate whether our results are robust to using alternative methods of textual analysis and to see whether the performance of the naïve Bayes model is dominated by the performance of alternative methods.

8.1 Weak Modal Words

The first method is based on the dictionary approach used in Ahern and Sosyura (2015), which is in turn based on a dictionary from Loughran and McDonald (2011) containing so-called weak modal words. Weak modal words such as “maybe,” “appears,” and “conceivable” have been shown by Ahern and Sosyura (2015) to indicate that merger rumors before the announcement are less likely to come true. It is therefore possible that in our setting, where we examine the announcement and post-announcement periods, weak modal words are also indicative of an approaching termination of the proposed merger.

To this end, we calculate for each press article the fraction of weak modal words contained in that article, using the weak modal word list from Ahern and Sosyura (2015). A larger fraction means that there appears more uncertainty in the article’s textual content, and we hypothesize that this uncertainty predicts a more likely termination of the merger.

In Figure 8, we repeat a similar analysis to the one we have previously conducted in Figure 3, the difference being that instead of plotting the naïve Bayes model, we plot the fraction of weak modal words. Specifically, we split up all mergers into two groups, one

Table XII. Can weak modal words predict merger outcomes?

This table shows a probit regression of a merger completion dummy variable on the percentage of weak modal words, which is an alternative measure of *ex ante* media content. Coefficients are multiplied by 100 for readability. Numbers in brackets show *t*-statistics. Stars indicate significance at 10%, 5%, and 1%.

	Ex post merger completion
Intercept (β_1)	1.28***
	(14.34)
Ex ante media content (β_2)	−2.66
	(−0.15)
McFadden R^2	0.00
Nagelkerke R^2	0.00
Number of observation	1,003

that will complete and one that will be withdrawn, and we plot the event-time score of weak modal words separately for both groups. We find that consistent with our hypothesis, withdrawn deals have a higher fraction of weak modal words. However, when contrasting Figure 8 with the earlier Figure 3, we demonstrate that weak modal words have a much smaller capacity than naïve Bayes to distinguish between deals that will complete and those that will be withdrawn.

To analyze the same question using a statistical test, we present in Table XII a similar assessment as previously displayed in Table IV, again with the difference that instead of naïve Bayes we use weak modal words to predict actual merger outcomes in a probit regression. Consistent with our hypothesis, we find that a higher fraction of weak modal words predicts merger withdrawal (i.e., we observe a negative coefficient since withdrawal is encoded as zero and completion as one). However, in contrast to naïve Bayes in Table IV, the coefficient is statistically insignificant and the R^2 is very close to zero.

When comparing weak modal words and naïve Bayes, we thus find that weak modal words have weaker capacity to accurately predict merger outcomes. This result is not entirely surprising, given that weak modal words are designed to capture "uncertainty," which is a less specific concept than "uncertainty about merger completion." In contrast, naïve Bayes is specifically trained to apprehend this latter notion by calculating the probability of merger completion (as described in Section 4), thus demonstrating improved ability to predict merger outcomes.

8.2 Tonal Words

An alternative to the weak modal words from the previous section is to look at tonal words. To this end, we use the positive and negative word lists from Loughran and McDonald (2011) and compute the fraction of positive to negative words in each press article. In analogy to our previous analysis of naïve Bayes and tonal words, we perform a basic sanity check to find out whether tonal words have predictive power for merger completion or failure. In Table XIII and Figure 9, we show the results of a probit regression of merger completion on tonal words and find that tonal words have predictive power. This result stands in contrast to weak modal words, where predictive power was absent. This outcome suggests that different words matter for different parts of the merger, that is, weak modal

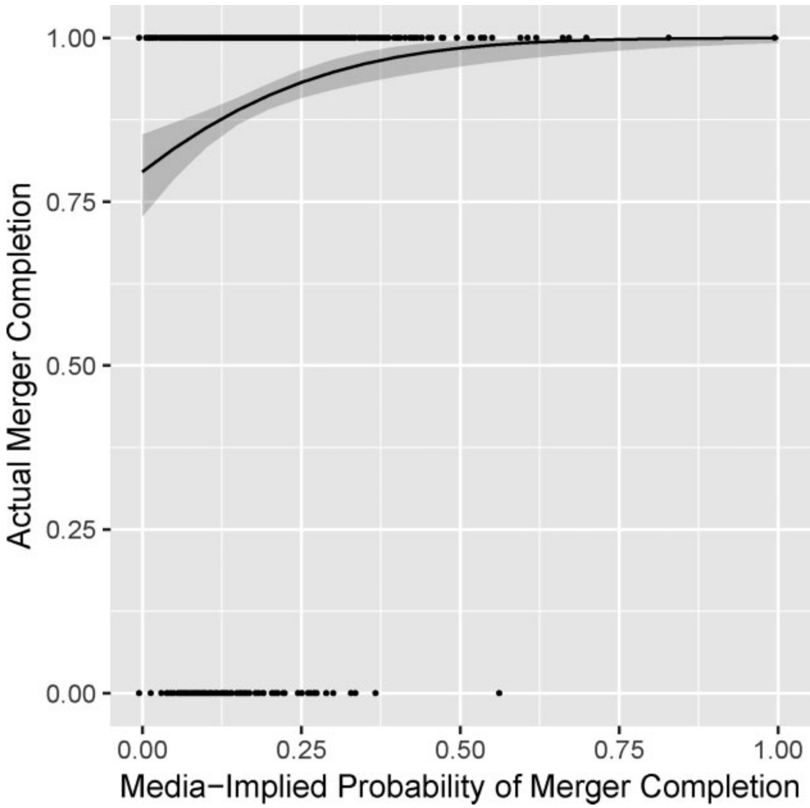

Figure 9. Media-implied probability of merger completion for tonal words. This plot shows the fitted line from a probit model and 95% confidence intervals, relating *ex ante* media content to *ex post* merger completion. The dependent variable on the *y*-axis is a dummy showing whether the merger completed in the end. The independent variable on the *x*-axis is the fraction of positive to negative words in press and newswire articles. The media measure is highly predictive because it uses only media information released on the announcement day, but not thereafter, while the dependent variable on the *y*-axis is measured at the merger's resolution date, which on average takes place more than one hundred days later.

words reflect rumors prior to the announcement as shown by Ahern and Sosyura (2015), while tonal words are related to merger completion after the announcement.

We then examine in Figure 10 whether tonal words can pick up differences between completed and withdrawn deals. The analysis is analogous to the one about naïve Bayes in Figure 3, where we split up the media measure into two subsets, one for deal that will complete and one for deals that will be withdrawn. The only difference is that we now use the media measure based on tonal words instead of naïve Bayes. We find that completed deals are associated on average with a higher occurrence of positive words, while withdrawn deals go together with an increase in negative words.

Given that completed deals are associated with more positive and fewer negative words, we study in Figure 11 whether trading on tonal words yields significant outperformance as measured by alpha. The analysis follows our previous tests from Figure 5. We invest in a merger deal if the media measure on the announcement day is above a given threshold level,

Table XIII. Tonal words predict merger outcomes

This table shows a probit regression of a merger completion dummy variable on the fraction of positive to negative words, which is an alternative measure of *ex ante* media content. Coefficients are multiplied by 100 for readability. Numbers in brackets show *t*-statistics. Stars indicate significance at 10%, 5%, and 1%.

	Ex post merger completion
Intercept (β_1)	0.83***
	(7.35)
Ex ante media content (β_2)	2.66***
	(4.19)
McFadden R^2	0.03
Nagelkerke R^2	0.04
Number of observation	1,000

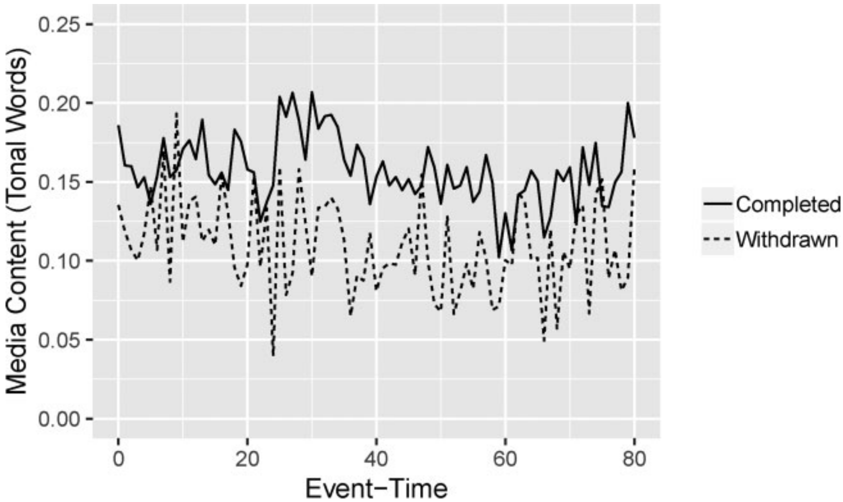

Figure 10. Tonal words in event time. This figure shows the average fraction of positive to negative words in event time of the merger. The solid line corresponds to announced mergers that will later complete, while the dotted line is for announced deals that will later be withdrawn.

shown on the x-axis. A higher threshold level means we exclude more deals with a higher failure probability according to the media measure. If markets react to the media measure, then a higher threshold level would correspond to a higher risk-adjusted return (alpha), thus resulting in an upward sloping line in Figure 11. The figure, however, fails to show consistently significant alphas, and furthermore indicates that if anything, the slope is downward sloping. (There is an upward jump on the right-hand side, which is highly insignificant and due to on outlier in the media measure, as can be seen from the sparsity of the upper dots on the right-hand side in Figure 9.) We therefore establish that markets do not react to tonal words. In comparison to the results of naïve Bayes (e.g., in Figure 5), we thus find that machine learning approaches such as naïve Bayes are necessary to extract information not yet incorporated in asset prices. Dictionary approaches such as weak modal words

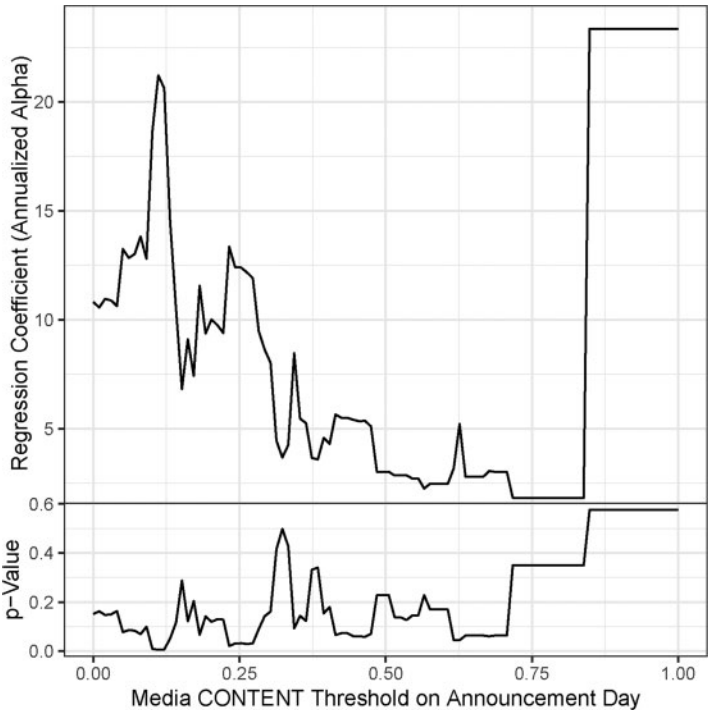

Figure 11. Simple trading strategy using tonal words. This figure shows risk-adjusted annualized returns ("alphas") from a simple trading strategy that conditions on media content on the announcement day, that is, the media-implied *ex ante* probability of merger completion. This plot is similar to the one in Figure 5 with the difference that instead of naïve Bayes we use the fraction of positive to negative words.

or tonal words are insufficient in the context of merger arbitrage to capture the richness of textual information contained in press and newswire articles.

8.3 Random Forests

Another model that has recently gained much traction in textual analysis applications is the random forests model from Ho (1995) and Breiman (2001). The basic idea behind random forests is to average several decision trees in order to overcome the well-known overfitting problem that plagues individual decision tress. In our context, a (very) stylized example of a decision tree would be a rule that predicts merger completion with, say, 95% if the words "poison pill" occur less than, say, two times in the press article. Otherwise (if "poison pill" occurs more than two times), the rule lowers the merger completion probability to, say, 30%. We use random forests as a plug-in replacement of naïve Bayes, that is, the function f in Equation (1).

Using random forests, we intend to uncover how naïve Bayes, which dates back at least to the early 1960s (Lewis, 1998), holds up in comparison to a model at the forefront of machine learning and computational linguistics. To this end, we show in Figure 12 the time series dynamics in event-time of the probability of merger completion constructed based on random forests. We draw two different lines, one relating to mergers that will later

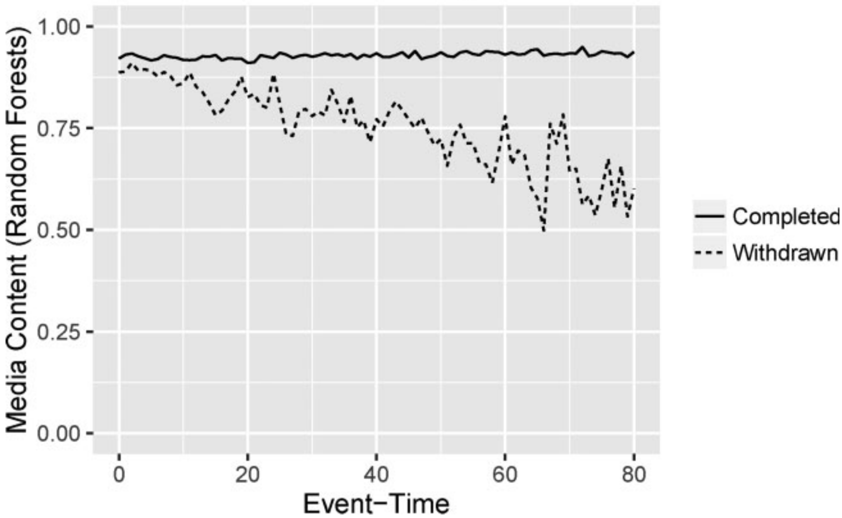

Figure 12. Random forests media measure in event-time. This figure shows the event-time dynamics of the random forests media content measure split up by whether the merger will complete in the end. Each line shows the media-implied probability of merger completion based on the random forests model. The plot is similar to the first one in Figure 3 with the difference that instead of naïve Bayes we use random forests for text classification.

Table XIV. Random forests predict merger completion

This table shows a probit regression of an *ex post* merger completion dummy on *ex ante* media content. In this table, media content is the probability of merger completion built up by the random forests model. This table is similar to Table IV, with the difference that instead of naïve Bayes we use random forests for text classification. Numbers in brackets show *z*-statistics. Stars indicate significance at 10%, 5%, and 1%.

	Ex post merger completion
Intercept (β_1)	-1.70^{***}
	(-2.60)
Ex ante media content (β_2)	3.28^{***}
	(4.53)
McFadden R^2	0.03
Nagelkerke R^2	0.04
Number of observation	960

complete and the other to those that will be withdrawn. We discover a clear and widening separation between both merger outcomes that expands as (event-) time goes by. We thus establish that random forests can isolate mergers that will complete from those that will fail, and that their forecast ability becomes more precise over time. Compared with Figure 3 however, which builds on naïve Bayes, the separation is less striking. Eighty days after a merger's announcement, random forests assign a success probability of around 50%

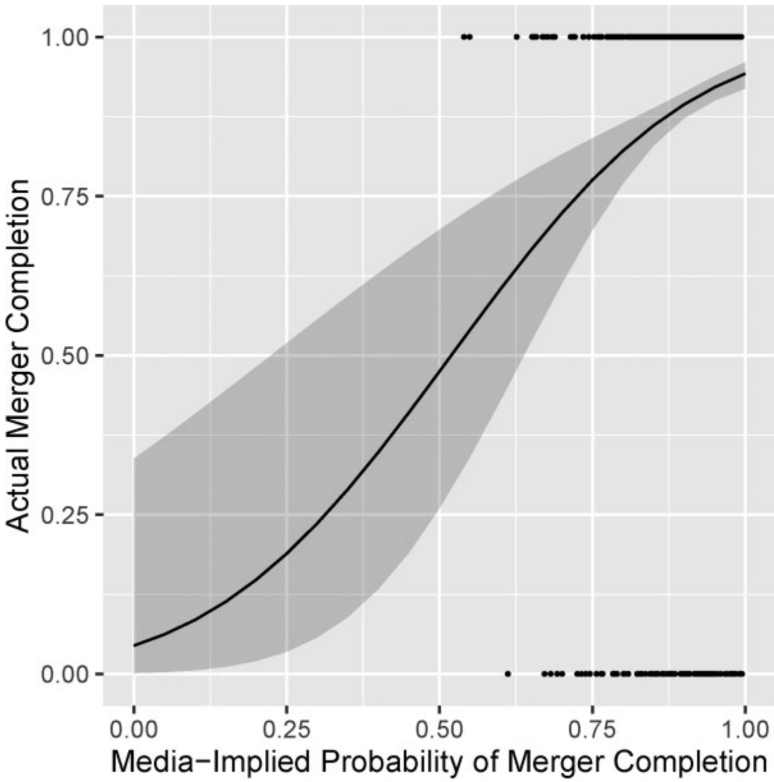

Figure 13. Media-implied probability of merger completion for random forests. This plot shows the fitted line from a probit model and 95% confidence intervals, relating *ex ante* media content to *ex post* merger completion. This plot is similar to the one in Figure 9 with the difference that instead of tonal words we use random forests for text classification.

to mergers that will later fail, while naïve Bayes assigns around 20%, showing that naïve Bayes apprehends the declining success probability more precisely.

In Table XIV, we run a similar probit regression as earlier in Table IV, with the difference that instead of naïve Bayes, we use random forests to calculate the media content measure (i.e., the media-implied probability of merger completion). We find that the significance of the media content coefficient as well as the R^2 is of similar magnitude, confirming that naïve Bayes and random forests perform equally well in predicting merger outcomes. The magnitude of the random forests coefficient is about 2.5 times as large as the naïve Bayes coefficient, which is due to the fact that random forests fail to capture the smaller probabilities of withdrawn deals (as noted in the discussion of Figure 12 above). Thus, the probit model compensates for the non-existing data points at the lower end by increasing the coefficient, as can be seen from the visualization of the probit model in Figure 13.

Since the random forests model is comparable in predictive performance to the naïve Bayes model, we examine next whether merger arbitrage returns react in a similar way to the media content measure based on random forests as they do to naïve Bayes. To this end, we repeat the investigation of the investment strategy from Section 6.2, where we exclude mergers that have a low *ex ante* completion probability. The difference is that instead of

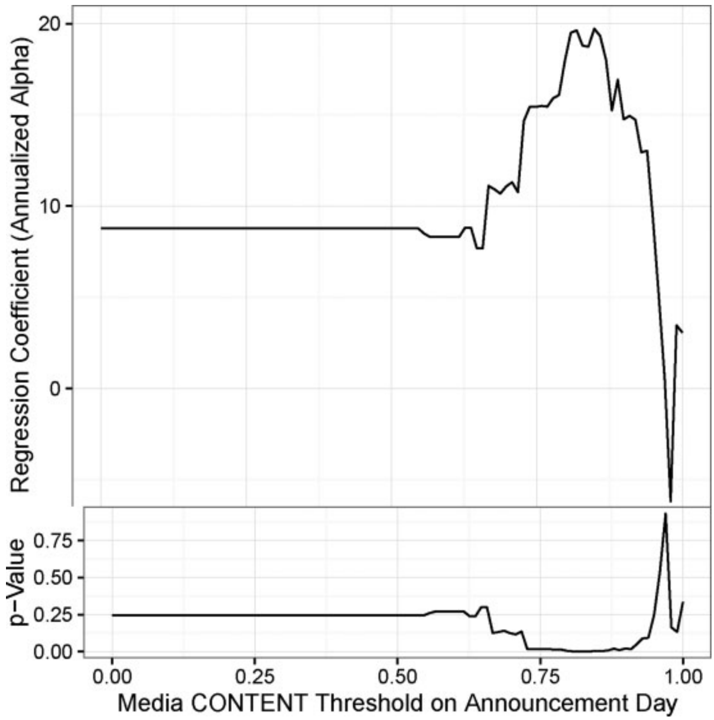

Figure 14. Simple trading strategy using random forests. This figure shows risk-adjusted annualized returns ("alphas") from a simple trading strategy that conditions on media content on the announcement day, that is, the media-implied *ex ante* probability of merger completion. This plot is similar to the one in Figure 5 with the difference that instead of naïve Bayes we use random forests for text classification.

naïve Bayes, we now use random forests to determine the completion probability. We hypothesize that the more mergers with low probability we eliminate, the higher should be the merger arbitrage return. To take this hypothesis to the data, we plot the risk-adjusted merger arbitrage returns (on the y-axis) as a function of a threshold level (on the x-axis) that removes all mergers with a completion probability below that threshold. To avoid look-ahead bias, we use media content from the announcement day (but not from thereafter) and start investing in a deal on the first trading day after the announcement (provided its above the threshold).

In Figure 14, we present a summary of this investigation. The first observation is that the line is flat for threshold levels below approximately 50%. This is due to the fact that, as noted above, random forests on our data do not generate probabilities below 50%. After we pass 50% however, Figure 14 demonstrates an initial increase in risk-adjusted merger arbitrage returns comparable to the naïve Bayes results shown earlier in Figure 5. The annualized returns increase to 20% from 8.8% as we continue to eliminate more low-probability mergers. However, as we keep increasing the threshold and remove a growing amount of low-probability mergers, the return performance starts to deteriorate. We thus establish that the random forests model yields comparable results to naïve Bayes as long as the threshold for removing low-probability mergers is not set too high.

Table XV. Regularized logit predicts merger completion

This table shows a probit regression of an *ex post* merger completion dummy on *ex ante* media content. In this table, media content is the probability of merger completion built up by the regularized logit model. This table is similar to Table IV, with the difference that instead of naïve Bayes we use regularized logit for text classification. Numbers in brackets show *z*-statistics. Stars indicate significance at 10%, 5%, and 1%.

	Ex post merger completion
Intercept (β_1)	0.29
	(1.17)
Ex ante media content (β_2)	1.08***
	(3.92)
McFadden R^2	0.02
Nagelkerke R^2	0.03
Number of observation	967

8.4 Regularized Logit

Another standard way to analyze textual content is the logit model, used for example by Routledge, Sacchetto, and Smith (2016) to predict whether a firm will be an acquirer or a target of an acquisition within 1 year of the firm's SEC filing. The logit model can be used as a plug-in replacement for naïve Bayes in our analysis, where the dependent variable in the regression is the merger's outcome (completion or withdrawal) and the independent variables are the word counts. Following Routledge, Sacchetto, and Smith (2016) we use a regularization technique known as "elastic net" to set some of the regression coefficients to zero if they do not contribute significantly to the predictive power of the model.

Words with the largest coefficients in absolute value include "unsolicited," "sweeten," "rebuff," "inadequate," "abandon," "reject," "prohibit," "unanimous," and "confidential." These words suggest that regularized logit identifies textual content that is economically meaningful within the context of the merger in question. For example, the chances of a merger deal succeeding or failing can be related to whether it is an "unsolicited" offer, whether the deal is "sweetened" (e.g., by a higher bid) or "rebuffed," whether the financing or bid price is "inadequate" or "rejected" and therefore potentially "abandoned." Likewise, a merger might have a high chance of success if there is "unanimous" agreement and when "confidential" negotiations have been held before.

In Table XV and Figure 15, we present a first sanity check to verify that the regularized logit model indeed does what it is supposed to do, that is, predict merger completion and failure. To this end, analogous to the previous tests with naïve Bayes (Table IV) and random forests (Table XIV and Figure 13), we run a probit regression of the merger completion dummy on the regularized logit media measure. We find that the media coefficient based on the regularized logit model is positive and highly significant, which shows that regularized logit does a good job at predicting merger outcomes.

We then repeat the remaining analysis of the regularized logit media measure in analogy to what we did before with robustness checks on random forests. Figure 16 demonstrates analogous to Figures 3 and 12 (naïve Bayes and regularized logit, respectively) that the media measure does a good job at picking up the increasing differences between completed

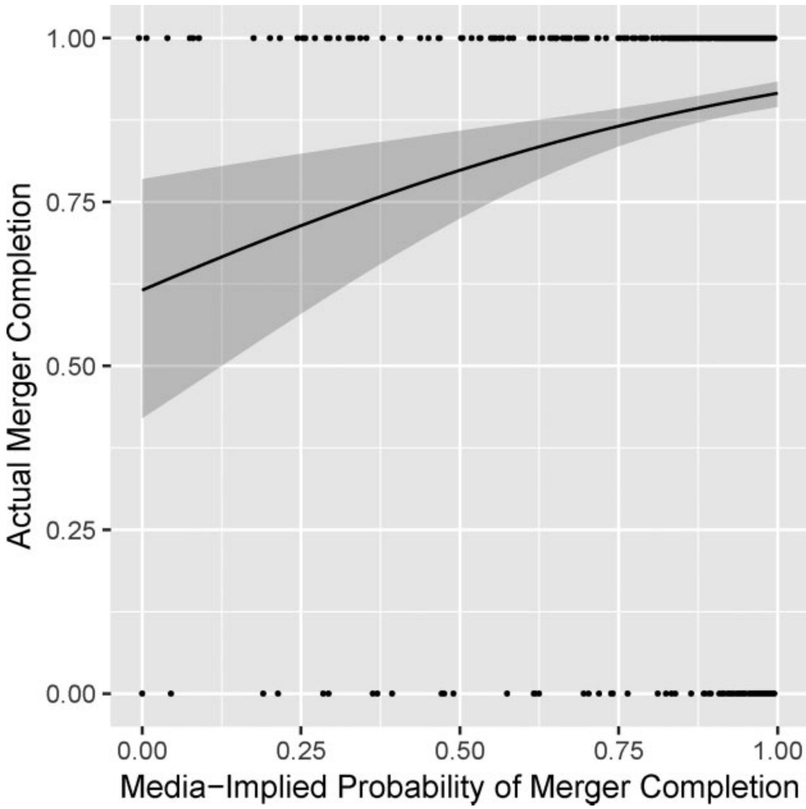

Figure 15. Media-implied probability of merger completion for regularized logit. This plot shows the fitted line from a probit model and 95% confidence intervals, relating *ex ante* media content to *ex post* merger completion. This plot is similar to the one in Figure 9 with the difference that instead of tonal words we use regularized logit for text classification.

and withdrawn merger deals as time progresses. And Figure 17 shows analogous to Figures 5 and 14 (naïve Bayes and regularized logit, respectively) that it is possible to trade on information from financial media if we quantify this information using regularized logit, as the alpha increases from about 10% to 20%. However, one important difference to naïve Bayes is that regularized logit has insignificant alphas over some threshold levels, notably for thresholds below 0.75 and above 0.95.

8.5 Extended Sample

In this section, we extend the sample to the beginning of 2016 to check whether our results stay robust. During this time, the use of computational methods to analyze text in newspapers and press releases has become more wide-spread. This effect could lead to more people trading on textual media content in mergers. Therefore, markets would become more efficient. If this is the case, the results would become attenuated or could even disappear.

Table XVI repeats the time series tests from our earlier Table VII with the extended sample. To take a conservative approach and to ensure that the power of the analysis is not mechanically driven by a larger sample size, we apply the same sample selection criteria as

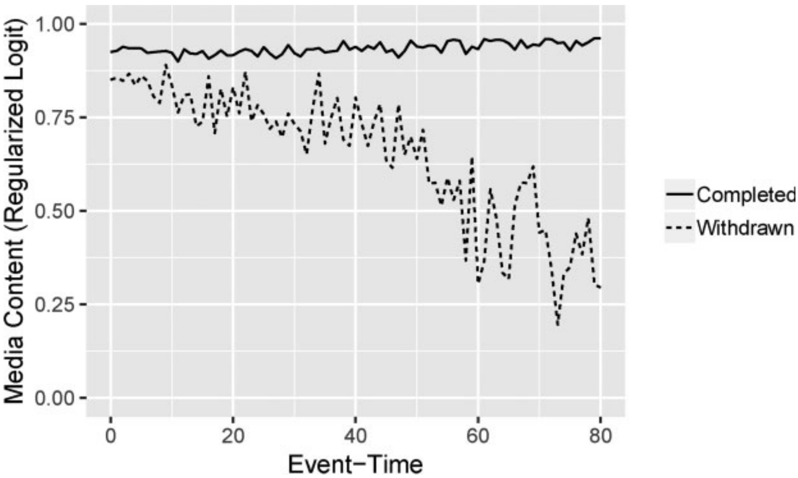

Figure 16. Regularized logit media measure in event-time. This figure shows the event-time dynamics of the regularized logit media content measure split up by whether the merger will complete in the end. Each line shows the media-implied probability of merger completion based on the regularized logit model. The plot is similar to the first one in Figure 3 with the difference that instead of naïve Bayes we use regularized logit for text classification.

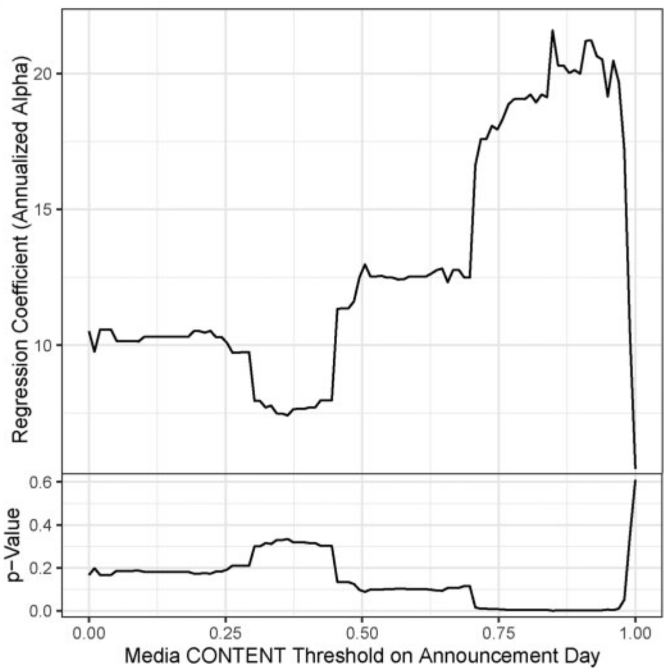

Figure 17. Simple trading strategy using regularized logit. This figure shows risk-adjusted annualized returns ("alphas") from a simple trading strategy that conditions on media content on the announcement day, that is, the media-implied *ex ante* probability of merger completion. This plot is similar to the one in Figure 5 with the difference that instead of naïve Bayes we use regularized logit for text classification.

Table XVI. Time series tests: extended sample

This table shows time series regressions of merger arbitrage portfolio returns on media measures. It is similar to Table VII, with the difference that an extended sample until the beginning of 2016 is used. Coefficients are multiplied by 100 for readability. Numbers in brackets show *t*-statistics. Stars indicate significance at 10%, 5%, and 1%.

	$r_{\mathrm{Tar}} - r_{\mathrm{f}}$	$r_{\mathrm{Tar}} - \delta r_{\mathrm{Acq}}$	$r_{\mathrm{Tar}} - r_{\mathrm{f}}$	$r_{\mathrm{Tar}} - \delta r_{\mathrm{Acq}}$
Intercept	−0.08	−0.08	−0.09	−0.11
	(−1.53)	(−1.16)	(−1.57)	(−1.50)
$r_{\mathrm{Mkt}} - r_{\mathrm{f}}$	71.38***	12.29***	71.38***	12.33***
	(80.33)	(10.94)	(80.32)	(10.98)
SMB	23.13***	4.14*	23.13***	4.09*
	(12.80)	(1.81)	(12.79)	(1.79)
HML	−19.77***	−5.25**	−19.78***	−5.35**
	(−11.85)	(−2.49)	(−11.85)	(−2.54)
Content lagged	0.15**	0.14*	0.15**	0.15*
	(2.28)	(1.77)	(2.29)	(1.80)
Coverage lagged			0.00	0.00**
			(0.37)	(2.13)
R^2	0.63	0.03	0.63	0.03
Number of observation	4,183	4,183	4,183	4,183

described in Section 5.1 to the extended time period. In particular, we start with the same number of deals (1,200) from SDC as before. We find in Table XVI that while the power decreases compared with the earlier sample, the results are still significant, particularly for the target's stock price reaction. Overall, we find that while the results become weaker, the media still has a significant effect on stock price reactions during mergers, most notably for the target.

9. Conclusion

Using merger announcements and applying methods from computational linguistics, we provide strong evidence that stock prices underreact to information in financial media. A one standard deviation increase in the media-implied probability of merger completion on the day of the announcement raises the return of a long-short merger strategy over the twelve post-announcement days by 1.2 percentage points. We find that the effect of a given increment in media-implied merger completion probability increases monotonically when we extend the holding period from one to 12 days after the announcement. Thus, media information takes several days before it is fully reflected in stock prices. In time-series analyses, we find that filtering out those deals with media-implied completion probabilities of 85% or less, which corresponds to 28% of the total sample, increases the annualized alpha from merger arbitrage by 9.3 percentage points.

The above findings vary significantly with financial market conditions. Following Axelson *et al.* (2013), we use the Merrill Lynch US High Yield Master II Option-Adjusted Spread to proxy for market conditions and find that media-based excess returns are particularly large and significant when it is hard for institutional investors to lever up, as indicated by a large high-yield spread. Specifically, annualized risk-adjusted returns increase by 11.3% when filtering out merger deals with low *ex ante* media-implied probability of merger completion. By contrast, such profits vanish when high-yield spreads are small.

We also analyze whether our findings are systematically related to the number of merger deals announced on a particular day. To this end we distinguish between Mondays when the number of announcements is typically large, and other days of the week. Indeed, we find that media information is only useful on days other than Mondays. Possibly, the information processing by financial media becomes noisier on days with many announcements.

While the main focus of the paper is the analysis of media content, that is, the particular words used in media and their information content for the completion probability, we also explore an alternative media measure based on media coverage. This measure only counts the number of press articles being released. While this is much simpler to measure, it may also be easier to manipulate (Ohl *et al.*, 1995; Ahern and Sosyura, 2014). Indeed, we find that the results based on this measure are much weaker. In fact, a trading strategy based on media coverage is statistically insignificant. Thus, to extract information about merger completion from financial media, it is important to analyze the content of those articles using textual analysis.

When investigating alternative procedures to analyze textual content, we demonstrate that dictionary approaches fail to fully capture important information about merger completion. On the other hand, we establish that random forests, an alternative model at the forefront of computational linguistics, exhibit a performance comparable to naïve Bayes when apprehending information about merger completion from textual content.

Taken together, our findings document that financial media contain fundamental information about the real economy that is not already reflected in stock prices. This study thereby contributes to our understanding of the limits and determinants of market efficiency, since our research design allows us to clearly identify the timing of information release and the subsequent market reactions. To this end, we provide a novel perspective on a specific channel of information transmission in financial markets and on the conditions which influence the speed at which this transmission takes place. Overall, we believe that methods of textual analysis together with the improved data availability have opened up a fruitful research path to improve our understanding of information aggregation in financial markets.

Supplementary Material

Supplementary data are available at *Review of Finance* online.

References

Ahern, K. R. and Sosyura, D. (2014): Who writes the news? Corporate press releases during merger negotiations, *Journal of Finance* 69, 241–291.

Ahern, K. R. and Sosyura, D. (2015): Rumor has it: sensationalism in financial media, *Review of Financial Studies* 28, 2050–2093.

Antweiler, W. and Frank, M. Z. (2004): Is all that talk just noise? The information content of internet stock message boards, *Journal of Finance* 59, 1259–1294.

Augustin, P., Brenner, M., and Subrahmanyam, M. G. (2019): Informed options trading prior to M&A announcements: insider trading?, *Management Science* 65, 5449–5956.

Axelson, U., Jenkinson, T., Strömberg, P., and Weisbach, M. S. (2013): Borrow cheap, buy high? The determinants of leverage and pricing in buyouts, *Journal of Finance* 68, 2223–2267.

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

Barber, B. M. and Odean, T. (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.

Barnes, B. G., Harp, L. N., and Oler, D. (2014): Evaluating the SDC mergers and acquisitions database, *Financial Review* 49, 793–822.

Beleites, C., Neugebauer, U., Bocklitz, T., Krafft, C., and Popp, J. (2013): Sample size planning for classification models, *Analytica Chimica Acta* 760, 25–33.

Bhattacharya, U., Galpin, N., Ray, R., and Yu, X. (2009): The role of the media in the internet IPO bubble, *Journal of Financial and Quantitative Analysis* 44, 657–682.

Bird, S., Klein, E., and Loper, E. (2009): *Natural Language Processing with Python*. O'Reilly Media, Sebastopol.

Brain, D. and Webb, G. I. (1999): On the effect of data set size on bias and variance in classification learning, in: *Proceedings of the Fourth Australian Knowledge Acquisition Workshop*, The University of New South Wales, Sydney, pp. 117–128.

Breiman, L. (2001): Random forests, *Machine Learning* 45, 5–32.

Buehlmaier, M. M. M. (2015): *The Role of the Media in Takeovers: Theory and Evidence*. Manuscript, University of Hong Kong.

Buehlmaier, M. M. M. and Whited, T. M. (2018): Are financial constraints priced? Evidence from textual analysis, *Review of Financial Studies* 31, 2693–2728.

Bushee, B. J. and Miller, G. S. (2012): Investor relations, firm visibility, and investor following, *Accounting Review* 87, 867–897.

Cao, C., Chen, Z., and Griffin, J. (2005): Informational content of option volume prior to takeovers, *Journal of Business* 78, 1073–1109.

Cao, C., Goldie, B. A., Liang, B., and Petrasek. L. (2016): What is the nature of hedge fund manager skills? Evidence from the risk arbitrage strategy, *Journal of Financial and Quantitative Analysis*, Forthcoming. 51, 929–957. doi:10.1017/S0022109016000387

Chan, K., Ge, L., and Lin, T.-C. (2013): Informational content of options trading on acquirer announcement return, *Journal of Financial and Quantitative Analysis*, Forthcoming 50, 1057–1082. doi:10.1017/S0022109015000484

Chan, W. S. (2003): Stock price reaction to news and no-news: drift and reversal after headlines, *Journal of Financial Economics* 70, 223–260.

Cook, D. O., Kieschnick, R., and Van Ness, R. A. (2006): On the marketing of IPOs, *Journal of Financial Economics* 82, 35–61.

Cornelli, F. and Li, D. D. (2002): Risk arbitrage in takeovers, *Review of Financial Studies* 15, 837–868.

Da, Z., Engelberg, J., and Gao, P. (2011): In search of attention, *Journal of Finance* 66, 1461–1499.

Da, Z., Gurun, U. G., and Warachka, M. (2014): Frog in the pan: continuous information and momentum, *Review of Financial Studies* 27, 2171–2218.

Das, S. R. and Chen, M. Y. (2007): Yahoo! for Amazon: sentiment extraction from small talk on the web, *Management Science* 53, 1375–1388.

Denis, D. J. and Macias, A. J. (2013): Material adverse change clauses and acquisition dynamics, *Journal of Financial and Quantitative Analysis* 48, 819–847.

Dinc, I. S. and Erel, I. (2013): Economic nationalism in mergers and acquisitions, *Journal of Finance* 68, 2471–2514.

Dougal, C., Engelberg, J., García, D., and Parsons, C. A. (2012): Journalists and the stock market, *Review of Financial Studies* 25, 639–679.

Duffie, D. (2010): Presidential address: asset price dynamics with slow-moving capital, *Journal of Finance* 65, 1237–1267.

Dukes, W. P., Frohlich, C. J., and Ma, C. K. (1992): Risk arbitrage in tender offers, *journal of Portfolio Management* 18, 47–55.

Dyck, A., Morse, A., and Zingales, L. (2010): Who blows the whistle on corporate fraud?, *Journal of Finance* 65, 2213–2253.

Dyck, A., Volchkova, N., and Zingales, L. (2008): The corporate governance role of the media: evidence from Russia, *Journal of Finance* 63, 1093–1135.

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

Engelberg, J. E., Reed, A. V. and Ringgenberg, M. C. (2012): How are shorts informed?: short sellers, news, and information processing, *Journal of Financial Economics* 105, 260–278.

Fama, E. F. and French, K. R. (1993): Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fang, L. and Peress, J. (2009): Media coverage and the cross-section of stock returns, *Journal of Finance* 64, 2023–2052.

Fang, L. H., Peress, J., and Zheng, L. (2014): Does media coverage of stocks affect mutual funds' trading and performance?, *Review of Financial Studies* 27, 3441–3466.

García, D. (2013): Sentiment during recessions, *Journal of Finance* 68, 1267–1300.

Giglio, S. and Shue, K. (2014): No news is news: do markets underreact to nothing?, *Review of Financial Studies* 27, 3389–3440.

Griffin, J. M., Hirschey, N. H., and Kelly, P. J. (2011): How important is the financial media in global markets?, *Review of Financial Studies* 24, 3941–3992.

Gurun, U. G. and Butler, A. W. (2012): Don't believe the hype: local media slant, local advertising, and firm value, *Journal of Finance* 67, 561–598.

Hanley, K. W. and Hoberg, G. (2010): The information content of IPO prospectuses, *Review of Financial Studies* 23, 2821–2864.

Hillert, A., Jacobs, H., and Müller, S. (2014): Media makes momentum, *Review of Financial Studies* 27, 3467–3501.

Hirshleifer, D. and Teoh, S. H. (2003): Limited attention, information disclosure, and financial reporting, *Journal of Accounting and Economics* 36, 337–386.

Ho, H. K. (1995): Random decision forests, in: *Proceedings of the Third International Conference on Document Analysis and Recognition*, Montreal, Quebec, Canada, 1995, pp. 278–282 vol.1, doi: 10.1109/ICDAR.1995.598994.

Hoberg, G. and Phillips, G. (2010): Product market synergies and competition in mergers and acquisitions: a text-based analysis, *Review of Financial Studies* 23, 3773–3811.

Hoberg, G., Phillips, G., and Prabhala, N. (2014): Product market threats, payouts, and financial flexibility, *Journal of Finance* 69, 293–324.

Huberman, G. and Regev, T. (2001): Contagious speculation and a cure for cancer: a nonevent that made stock prices soar, *Journal of Finance* 56, 387–396.

Jayaraman, N., Mandelker, G., and Shastri, K. (1991): Market anticipation of merger activities: an empirical test, *Managerial and Decision Economics* 12, 439–448.

Jegadeesh, N. and Wu, D. (2013): Word power: a new approach for content analysis, *Journal of Financial Economics* 110, 712–729.

Jensen, M. C. (1986): Agency costs of free cash flow, corporate finance, and takeovers, *American Economic Review* 76, 323–329.

Jetley, G. and Ji, X. (2010): The shrinking merger arbitrage spread: reasons and implications, *Financial Analysts Journal* 66, 54–68.

Jindra, J. and Walkling, R. A. (2004): Speculation spreads and the market pricing of proposed acquisitions, *Journal of Corporate Finance* 10, 495–526.

Kahneman, D. (1973): *Attention and Effort*. Prentice Hall.

Karolyi, A. and Shannon, J. (1999): Where's the risk in risk arbitrage? *Canadian Investment Review*, Spring. http://www.investmentreview.com/print-archives/spring-1999/features-departments-wheres-the-risk-in-risk-arbitrage-640/

Klibanoff, P., Lamont, O., and Wizman, T. A. (1998): Investor reaction to salient news in closed-end country funds, *Journal of Finance* 53, 673–699.

Kuhnen, C. M. and Niessen, A. (2012): Public opinion and executive compensation, *Management Science* 58, 1249–1272.

Larcker, D. F. and Lys, T. (1987): An empirical analysis of the incentives to engage in costly information acquisition: the case of risk arbitrage, *Journal of Financial Economics* 18, 111–126.

Lewis, D. D. (1998): Naive (Bayes) at forty: the independence assumption in information retrieval, in: C., Nédellec and C., Rouveirol (eds.), *Machine Learning: ECML-98, Number 1398 in Lecture Notes in Computer Science*, Springer, Berlin, pp. 4–15.

Löffler, G., Norden, L., and Rieber, A. (2019): *Salient News and the Stock Market Impact of Tone in Rating Reports*, Manuscript, Ulm University.

Li, F. (2010): The information content of forward-looking statements in corporate filings—a naïve Bayesian machine learning approach, *Journal of Accounting Research* 48, 1049–1102.

Liu, B. and McConnell, J. J. (2013): The role of the media in corporate governance: do the media influence managers' capital allocation decisions?, *Journal of Financial Economics* 110, 1–17.

Liu, L. X., Sherman, A. E., and Zhang, Y. (2014): The long-run role of the media: evidence from initial public offerings, *Management Science* 60, 1945–1964.

Lou, D. (2014): Attracting investor attention through advertising, *Review of Financial Studies* 27, 1797–1829.

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

Loughran, T. and McDonald, B. (2013): IPO first-day returns, offer price revisions, volatility, and form S-1 language, *Journal of Financial Economics* 109, 307–326.

Loughran, T. and McDonald, B. (2014a): Measuring readability in financial disclosures, *Journal of Finance* 69, 1643–1671.

Loughran, T. and McDonald, B. (2014b): Regulation and financial disclosure: the impact of plain English, *Journal of Regulatory Economics* 45, 94–113.

Lucca, D. O. and Trebbi, F. (2011): *Measuring Central Bank Communication: An Automated Approach with Application to FOMC Statements*, Manuscript, University of British Columbia.

Manela, A. (2014): The value of diffusing information, *Journal of Financial Economics* 111, 181–199.

Manning, C. D., Raghavan, P., and Schütze, H. (2008): *Introduction to Information Retrieval*, 1st edition, Cambridge University Press, New York.

Manning, C. D. and Schütze, H. (1999): *Foundations of Statistical Natural Language Processing*, MIT Press, Cambridge, MA.

Mitchell, M., Pedersen, L. H., and Pulvino, T. (2007): Slow moving capital, *American Economic Review* 97, 215–220.

Mitchell, M. and Pulvino, T. (2001): Characteristics of risk and return in risk arbitrage, *Journal of Finance* 56, 2135–2175.

Mitchell, M. and Pulvino, T. (2012): Arbitrage crashes and the speed of capital, *Journal of Financial Economics* 104, 469–490.

Officer, M. S. (2007): Are performance based arbitrage effects detectable? Evidence from merger arbitrage, *Journal of Corporate Finance* 13, 793–812.

Ohl, C. M., Pincus, J., Rimmer, T., and Harrison, D. (1995): Agenda building role of news releases in corporate takeovers, *Public Relations Review* 21, 89–101.

Palepu, K. G. (1986): Predicting takeover targets, *Journal of Accounting and Economics* 8, 3–35.

Peress, J. (2014): The media and the diffusion of information in financial markets: evidence from newspaper strikes, *Journal of Finance* 69, 2007–2043.

Porter, M. F. (1980): An algorithm for suffix stripping, *Program* 14, 130–137.

Rish, I. (2001): An empirical study of the naive bayes classifier, in: *Proceedings of IJCAI-2001 workshop on Empirical Methods in AI*, vol. 3, Seattle, pp. 41–46.

Romer, C. D. and Romer, D. H. (2004): A new measure of monetary shocks: derivation and implications, *American Economic Review* 94, 1055–1084.

Rosa, C. (2011): The high-frequency response of exchange rates to monetary policy actions and statements, *Journal of Banking and Finance* 35, 478–489.

Rosa, C. and Verga, G. (2007): On the consistency and effectiveness of central bank communication: evidence from the ECB, *European Journal of Political Economy* 23, 146–175.

Routledge, B. R., Sacchetto, S., and Smith, N. A. (2016): *Predicting Merger Targets and Acquirers from Text*, Manuscript, Carnegie Mellon University.

Scherbina, A. and Schlusche, B. (2015): *Economic linkages inferred from news stories and the predictability of stock returns*, Manuscript, University of California, Davis.

Schmeling, M. and Wagner, C. (2019): *Does Central Bank Tone Move Asset Prices*? Manuscript, City University London.

Solomon, D. H. (2012): Selective publicity and stock prices, *Journal of Finance* 67, 599–638.

Solomon, D. H., Soltes, E., and Sosyura, D. (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, 1139–1168.

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

Tetlock, P. C. (2011): All the news that's fit to reprint: do investors react to stale information?, *Review of Financial Studies* 24, 1481–1512.

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

Veldkamp, L. L. (2006): Media frenzies in markets for financial information, *American Economic Review* 96, 577–601.

Wang, B. (2017): *Probability Weighting and Asset Prices: Evidence from Mergers and Acquisitions*, Manuscript, Fordham University.

Zhang, H. (2005): Exploring conditions for the optimality of naïve bayes, *International Journal of Pattern Recognition and Artificial Intelligence* 19, 183–198.