Takeovers and the Media

Matthias Buehlmaier*

December 7, 2009

Abstract

This paper develops and empirically confirms a theory that explains why media content predicts takeover outcomes. It shows why target shareholders pay attention to the news, despite the risk of distorted reporting. To test the model's prediction, this paper constructs a novel empirical measure that quantifies text-based media content pertaining to the acquirer. It then estimates the intertemporal link between this measure and the outcome of takeovers. Consistent with theory, positive media content about the acquirer predicts takeover success. Relative to other predictors proposed in the literature, the media measure is the most important explanatory variable in terms of significance and goodness of fit. Using text-based media content, this paper thus fills a major gap in our understanding of takeovers.

*Vienna Graduate School of Finance (VGSF), Heiligenstaedterstr. 46–48, 1190 Vienna, Austria. Phone: +43 (1) 31336–5992, email: matthias.buehlmaier@vgsf.ac.at, web: www.buehlmaier.net. My greatest debt is to my advisors Klaus Ritzberger and Josef Zechner, who provided me with invaluable advice and comments. I would like to express my deeply felt gratitude to them as they have constantly and generously shared their knowledge with me. I would like to thank Filippo Balestrieri, Jonathan Berk, Gerhard Binas, Thomas Dangl, Stefano Demichelis, Andrea Eisfeldt, Alexander Eisl, Vito Gala, Alois Geyer, Hamed Ghoddusi, Klaus Gugler, Christian Haefke, Christopher Hennessy, Emir Hrnjic, Philipp Illeditsch, Natalia Ivanova, Michael Kisser, Mathieu Luypaert, Fabio Maccheroni, Kristian Miltersen, Alexander Muermann, David Musto, Timofiy Mylovanov, Stijn Van Nieuwerburgh, Terrance Odean, Gordon Phillips, Pegaret Pichler, Ehud Ronn, Johannes Ruf, Jacob Sagi, Stephen Shore, Jamsheed Shorish, Baran Siyahhan, Paolo Sodini, Leopold Sögner, Fallaw Sowell, Helmut Strasser, Oren Sussman, Sheridan Titman, Laura Veldkamp, Alexander Wagner, Neng Wang, Jörgen Weibull, Russ Wermers, Toni Whited, Youchang Wu, Wei Xiong, Vilimir Yordanov, Alexander Zimper, Jeffrey Zwiebel, and the participants in my presentations at Vienna University of Economics and Business for valuable discussions and comments. I am very grateful for the generous support of VGSF.

Language is only the instrument of science, and words are but the signs of ideas: I wish, however, that the instrument might be less apt to decay, and that signs might be permanent, like the things they denote.

Samuel Johnson (Preface, *A Dictionary of the English Language*, April 15, 1755)

Anecdotal evidence suggests that the financial media plays a key role in mergers and acquisitions. Yet, it is unclear why media content goes beyond public information available elsewhere. This paper characterizes the relationship between the acquirer, the media, and target shareholders, focusing on the immediate influence of media content on takeover outcome.

To my knowledge, this paper is the first to develop and empirically confirm a theory that explains why media reports predict takeover outcomes. Endogeneity resulting from the strategic interaction of agents requires a game-theoretic approach to derive the following empirical prediction: positive information in the media about the acquirer entails takeover success. This prediction obtains because the media endogenously mitigates the informational asymmetry between the acquirer and target shareholders, who decide on the takeover outcome. Using naïve Bayes, a simple and well-known probabilistic model, this paper constructs an empirical measure indicating how positive the media is about the acquirer. It then estimates the intertemporal link between this measure and the probability of takeover success and confirms[1] the empirical prediction.

The economic intuition follows from two key considerations. First, target shareholders do not know whether a takeover leads to the creation or destruction of target value. Since target shareholders know their own company, their uncertainty pertains to the acquirer's type. All types have incentives to take over the target. The good type gains because it directly profits from the value creation after a successful takeover. The bad type gains because of private benefits of control. Since shareholders do not know the acquirer's type, they are uncertain whether or not to approve[2] the takeover.

Second, the acquirer has the opportunity to run a costly media campaign. This consideration is motivated by Ohl et al. (1995), who show in a takeover context that companies' press releases and media access to company executives influence media content. By reading between the lines, shareholders obtain information on whether or not a media campaign occurred. Since some residual uncertainty remains, shareholders receive a noisy signal. In equilibrium, a signal indicating the occurrence or absence of a media campaign carries information about the acquirer. It thus mitigates the informational asymmetry between acquirer and target shareholders. The signal carries information because it is optimal for the good and the bad type to play different media strategies. The good type has incentives to release as much information as possible about

[1] That is, the sample data does not reject this paper's hypothesis.

[2] Approval takes place in the form of a shareholder vote or shareholders' tendering decision. This paper's theoretical part focuses on the more challenging case of shareholders' tendering decision. This case is more challenging because of potential free-rider problems. The case of shareholder voting trivially obtains by replacing the tendering decision with an arbitrary voting system. The media dynamics and the empirical prediction stay the same.

his type: only if target shareholders are sufficiently convinced that the good type is behind the media campaign they do approve the takeover. The good type thus runs a media campaign to distinguish itself from the low type. The bad type abstains from a media campaign because it is too costly to mimic the good type.

Using naïve Bayes to construct the empirical media measure is a natural choice. Defining in general what constitutes a news article that is positive about the acquirer is challenging. It is challenging because text-based information in a takeover context is multilayered and nontrivial. Counting the number of positive words, for example, might not suffice: it matters whether or not positive words pertain to the acquirer. The solution to this challenge is in the spirit of Potter Stewart:[3] positive text-based information about the acquirer is hard to define, "but I know it when I see it." I obtain a small[4] estimation sample by manually classifying news articles into two groups consisting of articles that are positive and negative about the acquirer. Antweiler and Frank (2004) show that naïve Bayes works very well with unstructured text and a similar classification scheme. I then apply the estimated naïve Bayes model to the large number of remaining news articles out-of-sample. The empirical media measure obtains by the aggregation of media content pertaining to a given takeover.

The takeover attempt's outcome is either success or failure and thus a binary variable. This paper tests the empirical prediction using standard specifications of a general binary outcome model. The media measure predicts the outcome of takeovers statistically significant and economically meaningful. The inclusion of the media measure drastically increases goodness of fit. Considering acquirer and target firm characteristics, only acquirer characteristics are significant. This result confirms the game-theoretic model's assumption that shareholder uncertainty pertains to the acquirer only. Robustness checks are obtained by considering a different aggregation of text-based information and by using an alternative binary outcome model.

This paper explicitly addresses three empirical sources of potential bias. First, it solves endogeneity of the media measure by excluding news articles that appear on the effective date or the date withdrawn. The media measure is thus exogenous with respect to takeover outcome. Second, this paper avoids data snooping bias because of the game-theoretic development of its empirical prediction and because of the out-of-sample construction of the media measure. Third, it avoids sample selection bias due to right censoring. It considers a time period implying a sample free of duration-induced bias because the outcome of every announced takeover attempt is known.

Section 1 presents the game-theoretic model. Section 2 solves the model and derives the resulting equilibrium. Section 3 states the empirical prediction. Section 4 introduces the data. Section 5 shows the empirical methodology. It details how naïve Bayes is used to construct the media measure and introduces the binary outcome model. Section 6 shows estimation results and Section 7 contains robustness checks. Section 8 concludes and the Appendix contains the proofs pertaining to Section 2.

[3]Potter Stewart (January 23, 1915 – December 7, 1985) was an Associate Justice of the United States Supreme Court.

[4]The estimation sample size is less than 0.5% of the total sample size.

Related Literature Antweiler and Frank (2004) are among the first to address the interpretation of unstructured text in a financial context. They study the effect of messages posted on Yahoo! Finance and Raging Bull on two stock market indices. As this paper, Antweiler and Frank use naïve Bayes to quantify text-based information. Confirming Harris and Raviv (1993), they find that disagreement among posted messages is associated with increased trading volume. Although stock messages have a statistically significant effect on stock returns, their economic significance is small.

Tetlock (2007) studies the effect of a popular Wall Street Journal column on the stock market. He uses an automated program, the General Inquirer, to analyze the column. The General Inquirer counts words that fall within 77 categories from the Harvard psychosocial dictionary. Using principal component analysis, Tetlock identifies the most relevant categories and derives a pessimism factor. He shows that the pessimism factor predicts downward pressure on stock prices followed by a reversion and that high or low pessimism predicts high trading volume. This finding is consistent with theoretical results on noise and liquidity traders.

Veldkamp (2006) relates surges in prices and cross-market price dispersion to media coverage. Veldkamp explains these anomalies with an information market complementarity. Her result obtains because information is fundamentally distinct from other goods. It has a fixed cost of discovery and a near-zero cost of replication. High volume thus makes information inexpensive and low prices induce investors to buy information that others also buy. Veldkamp finds empirical support for the model's prediction that asset market movements generate news and that news raise prices and price dispersion.

1 The Model

The following game is in the spirit of the classical[5] Bagnoli and Lipman (1988) model. Figure 1 illustrates the game's timeline. There are two companies, the acquirer and the target. The acquirer is either of high H or of low L type. The high type increases the value of the target in case of a takeover while the low type destroys value. A takeover thus causes the value of a share of target stock to change to p_t from p, where $t \in \{H, L\}$ and $p_L < p < p_H$. This assumption is motivated by standard arguments pertaining to potential benefits and costs of takeovers. Examples include agency costs, managerial ability, efficiency, operating and financial synergy, taxes, and market power.

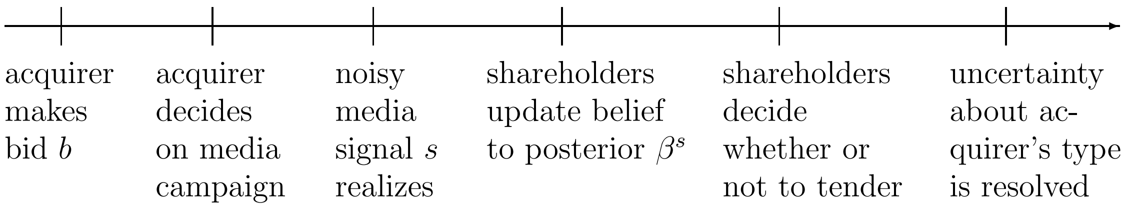
The n target shareholders do not know the acquirer's type. They share a common prior

$$\beta := P(T = H) \in (0, 1),$$

where T denotes a random variable with realizations $t \in \{H, L\}$ representing target shareholders' uncertainty about the acquirer's type.

[5] Bagnoli and Lipman (1988) overcome the free-rider problem of Grossman and Hart (1980) by making shareholders pivotal.

Figure 1: Timeline



At the initial stage the acquirer decides on the bid price $b \geq 0$ of its any-and-all bid. After making the bid, the acquirer obtains previously unknown information about the target through interaction with target shareholders or target management. The acquirer learns how a potential merger would affect the target's value. The acquirer thus privately learns its type t. Knowing its type, the acquirer decides whether or not to run a media campaign: the acquirer decides on its potentially degenerate mixed strategy[6]

$$\xi_t := P(M = 1 | \, T = t) \in [0, 1], \qquad t \in \{H, L\}, \tag{1}$$

where the event $\{M = 1\}$ denotes the occurrence of a media campaign and $\{M = 0\}$ denotes its absence. A media campaign is costly. The acquirer thus incurs costs $c > 0$ if the event $\{M = 1\}$ occurs, that is, if the realization m of the random variable M is equal to one.

Target shareholders cannot observe the realization m, that is, they do not know whether or not the acquirer engaged in a media campaign. Reading between the lines, they instead obtain a noisy signal $s \in \{0, 1\}$. The signal s is a realization of the random variable S, where S is positively correlated with the media campaign. The signal precision

$$\delta := P(S = 1 | \, M = 1) = P(S = 0 | \, M = 0) > \frac{1}{2}$$

determines the conditional dependence of S on M. Shareholders incorporate the new information contained in the signal by updating their belief to the posterior

$$\beta^s := P(T = H | \, S = s).$$

The primary reason for shareholder uncertainty about the acquirer's type is uncertainty about the acquirer's actions implied by ξ_H and ξ_L. Although the noisy signal makes this model more realistic, its implications for shareholder uncertainty are of second order: even $\delta = 1$ does not resolve uncertainty if at least one acquirer type plays a non-degenerate mixed strategy.

At the end of the game, target shareholders decide whether or not to tender their shares. Each shareholder holds one share.[7] If shareholders tender at least k shares, the takeover succeeds.

[6] The acquirer chooses a random variable M with realizations in $\{0, 1\}$. Only the distribution of M matters in this game. One can thus think of ξ_t as a choice variable W.L.O.G.

[7] This simplifying assumption is weak: the equilibrium results stay unchanged if shareholders hold more than one share and if different shareholders hold different amounts of shares.

Otherwise it fails. If a shareholder does not tender and the takeover succeeds, his expected payoff is $p^s - p$, where

$$p^s := \beta^s p_H + (1-\beta^s)p_L$$

is shareholders' expected post-takeover value per target share. If a shareholder tenders and the takeover succeeds, his payoff is $b - p$. If the takeover fails, each shareholder's payoff is zero.

Acquirer management obtains nonmonetary private benefits of control $z > k(p - p_L)$ from running the target. The lower bound on z ensures that the low type has incentives to bid for the target although it reduces target value. Acquirer management's final payoff is

$$-c\mathbb{1}_{m=1} + [z + j(p_t - b)]\mathbb{1}_{j \geq k},$$

where j is the number of shares tendered and $\mathbb{1}$ is the indicator function.

2 Equilibrium

Consider target shareholders' decision to tender. Distinguish between the following four cases. First, suppose[8] that $p \vee p^s < b$. All shareholders have incentives to tender. The takeover thus succeeds and the acquirer obtains $-c\mathbb{1}_{m=1} + z + n(p_t - b)$. Second, suppose that $b < p \wedge p^s$. No shareholder has incentives to tender. The takeover thus fails and the acquirer obtains $-c\mathbb{1}_{m=1}$. Third, suppose that $p < b < p^s$. Focus on pure tendering strategies. Then any set of shareholder strategies in which exactly k shares are tendered is an equilibrium. Those shareholders that do tender have no incentive to deviate: if they would not tender, the takeover would fail and they would receive p instead of b. Those shareholders that do not tender also have no incentive to deviate: if they would tender, the takeover would still succeed and they would receive b instead of p^s. As a result, exactly k shares are tendered, the takeover succeeds, and the acquirer obtains $-c\mathbb{1}_{m=1} + z + k(p_t - b)$.

Fourth, suppose that $p^s < b < p$. One equilibrium is that no shareholder tenders. No shareholder has incentives to deviate: if a shareholder would tender, the takeover would still fail and his utility would stay unchanged. The takeover thus fails and the acquirer obtains $-c\mathbb{1}_{m=1}$. Another equilibrium is that every shareholder tenders. No shareholder has incentives to deviate: if a shareholder would not tender, the takeover would still succeed and the shareholder would obtain p^s instead of b. This equilibrium is in the spirit of the bank run equilibrium of Diamond and Dybvig (1983). It gives the acquirer perverse incentives to convince shareholders that it is unfit to run the target company. This equilibrium is implausible because it would lead to white knight counterbids or to regulatory intervention. I do not focus on this equilibrium in the remainder of this paper.

Target shareholders' tendering strategies in case of shareholder indifference need to ensure that the acquirer's maximization problem with respect to b, ξ_H, and ξ_L is well-behaved. Otherwise no equilibrium exists. Thus, target shareholders tender k shares if $b = p \leq p^s$ or if $b = p^s > p$.

[8] This paper uses the notation $a \vee b = \max\{a, b\}$ and $a \wedge b = \min\{a, b\}$.

Lemma 1 (Optimal Bid)**.** *The acquirer bids $b^* = p$ in equilibrium.*

If the acquirer bids less than p, no shareholder tenders and the takeover fails with probability one. If the acquirer bids $b > p$, the takeover succeeds. Halving the spread between b and p implies that the acquirer increases its expected payoff: it pays less and the takeover still succeeds. The acquirer thus bids $b^* = p$ in equilibrium.

Lemma 1 shows that the acquirer obtains at most the amount $z + k(p_H - p)$ from a successful takeover. To make the model interesting, the cost of the media campaign must be lower than the maximal amount gained in the takeover. Otherwise an empty set of nontrivial equilibria obtains because no acquirer type has incentives to run a media campaign. Therefore, suppose that

$$c < [z + k(p_H - p)](2\delta - 1) =: \bar{c}. \tag{2}$$

holds for the remainder of this paper.

The following lemma shows how target shareholders incorporate the new information obtained from the noisy signal s into their posterior beliefs β^s.

Lemma 2 (Posterior Beliefs)**.** *Target shareholders' posterior beliefs are*

$$\begin{aligned} \beta^0 &= \beta \cdot \frac{(1-\delta)\xi_H + \delta(1-\xi_H)}{1-\zeta}, \\ \beta^1 &= \beta \cdot \frac{\delta\xi_H + (1-\delta)(1-\xi_H)}{\zeta}, \end{aligned} \tag{3}$$

where $\zeta = P(S = 1) = \delta\mu + (1-\delta)(1-\mu)$ and $\mu = P(M = 1) = \xi_H\beta + \xi_L(1-\beta)$.

The economic intuition follows from the observation that

$$\beta^s_{\xi_H} \begin{cases} < 0 & , s = 0 \\ > 0 & , s = 1 \end{cases}, \qquad \beta^s_{\xi_L} \begin{cases} > 0 & , s = 0 \\ < 0 & , s = 1 \end{cases},$$

where $\beta^s_{\xi_t}$ denotes the partial derivative with respect to ξ_t. Consider for example the implications of $\beta^1_{\xi_H} > 0$. Shareholders observe the signal realization $s = 1$. Since the signal is positively correlated with the media campaign, shareholders believe that a media campaign occurred with high probability. First, suppose the high acquirer type runs a media campaign with a low probability. That is, ξ_H is close to zero. Since a media campaign occurred with high probability and since shareholders know that the high type is unlikely to run a media campaign, they are relatively certain that they face a low type. That is, β^1 is relatively small. Second, suppose the high type runs a media campaign with high probability. That is, ξ_H is close to one. Since a media campaign occurred with high probability and since shareholders know that the high type is very likely to run a media campaign, they are relatively certain that they face a high type. That is, β^1 is relatively large. This argument illustrates why the posterior β^1 is an increasing function of ξ_H.

Define

$$\underline{\beta} := \frac{(p - p_L)(1 - \delta)}{(p - p_L)(1 - \delta) + (p_H - p)\delta},$$
$$\bar{\beta} := \frac{(p - p_L)\delta}{(p - p_L)\delta + (p_H - p)(1 - \delta)}$$

and let $\underline{c} := [z - k(p - p_L)](2\delta - 1)$.

Lemma 3. *It holds that* $0 < \underline{\beta} < \bar{\beta} < 1$ *and that* $\underline{c} < \bar{c}$.

Lemma 3 shows how $\underline{\beta}$ and $\bar{\beta}$ partition the unit interval into the three subintervals $[0, \underline{\beta})$, $[\underline{\beta}, \bar{\beta})$, and $[\bar{\beta}, 1]$. The economic intuition pertains to shareholders' prior belief about the acquirer. If $\beta \in [0, \underline{\beta})$, shareholders are relatively pessimistic about the acquirer because they are relatively certain that they face a low type. If $\beta \in [\underline{\beta}, \bar{\beta})$, shareholder uncertainty about the acquirer's type is relatively large. Neither do they believe that they face a high nor a low type with relatively high certainty. Finally, if $\beta \in [\bar{\beta}, 1]$, shareholders are relatively certain to be facing a high type.

Theorem (Perfect Bayesian Equilibrium)**.**

- *If* $\beta \in [\underline{\beta}, \bar{\beta})$ *and* $c \geq \underline{c}$, *then* $(\xi_H^*, \xi_L^*) = (1, 0)$. *The takeover succeeds after* $s = 1$ *and fails after* $s = 0$.
- *If* $\beta \notin [\underline{\beta}, \bar{\beta})$ *or* $c < \underline{c}$, *then* $(\xi_H^*, \xi_L^*) = (0, 0)$. *The takeover succeeds if* $\beta \in [\underline{\beta}, 1]$ *and it fails otherwise.*

First, consider the separating equilibrium $(\xi_H^*, \xi_L^*) = (1, 0)$. Shareholder uncertainty about the acquirer is large ($\beta \in [\underline{\beta}, \bar{\beta})$). The costs of running a media campaign are intermediate ($c \in [\underline{c}, \bar{c}]$). The high type knows that it is going to profit from the post-takeover increase in target value even if it pays for the media campaign. The high type has thus incentives to distinguish itself from the low type by running a media campaign. The low type has no incentives to mimic the high type because a media campaign is too costly. It is too costly because the low type's private benefits of control do not suffice to compensate for both the decrease in target value and the costs of the media campaign. Since the media signal is positively correlated with the media campaign, shareholders believe that a takeover creates value only if they observe a positive media signal $s = 1$. Since shareholder beliefs determine the takeover's outcome, a positive media signal entails takeover success and a negative signal entails failure.

Second, suppose that a media campaign is relatively cheap ($c < \underline{c}$). Because the media campaign is cheap, the low type has incentives to mimic the high type. The high type is unable to distinguish itself from the low type. The high type thus has no incentive to spend money on the media campaign. A pooling equilibrium obtains. If shareholders are sufficiently optimistic about the acquirer ($\beta \in [\underline{\beta}, 1]$), the takeover succeeds and fails otherwise.

Third, suppose target shareholders are relatively certain that the acquirer is of high type ($\beta \in [\bar{\beta}, 1]$). Shareholders are so certain that a takeover will improve target value that they tender k shares even if they do not receive additional information (i.e. if $\beta^s = \beta$). No acquirer type has thus incentives to release information by spending money on a media campaign to distinguish itself from the other type. Both acquirer types abstain from a media campaign $((\xi_H^*, \xi_L^*) = (0, 0))$ and the takeover succeeds with probability one.

Fourth, suppose that target shareholders are relatively certain that the acquirer is of low type ($\beta \in [0, \underline{\beta})$). Shareholders are so pessimistic about the acquirer's type that no media strategy can change their opinion. Knowing that the takeover will not succeed, both types abstain from paying for a media campaign.

3 Empirical Prediction

The above theorem shows that even under weak assumptions the media plays a key role in mergers and acquisitions. The only conditions for the existence of a nontrivial media equilibrium are (i) there is sufficient shareholder uncertainty about the acquirer ($\beta \in [\underline{\beta}, \bar{\beta})$) and (ii) the cost of running a media campaign is intermediate ($c \in [\underline{c}, \bar{c}]$). These conditions hold for almost all takeovers.

The theorem shows that a positive media signal carries positive information about the acquirer and thus entails takeover success. This consideration immediately yields the following empirical prediction:

> Positive information in the financial media about the acquirer increases the probability of takeover success.

4 Data

Takeover data are from the SDC Platinum Mergers & Acquisitions database pertaining to US targets. The criteria for sample inclusion are as follows. Only takeover attempts with announcement dates between January 1, 2000 and December 31, 2006 are included. This choice of time period is motivated by two considerations. First, it avoids sample selection bias due to right censoring. If more recent announcement dates would be included, the sample would have to exclude takeovers with yet unknown outcome. This exclusion leads to bias because it is nonrandom. Second, the chosen time period covers approximately one business cycle. This consideration matters because of merger waves.

The sample excludes the industries "energy and power," "financials," and "government and agencies." Only takeover attempts with enterprise value at announcement of at least \$500 million are included. The motivation for this criterion is that media coverage mainly pertains to large deals. By a similar argument, the sample only includes publicly traded acquirers and targets. A second reason for this inclusion criterion is that the media is of particular importance for widely-held firms. Cross-border deals are excluded: government policy, regulation, tariffs, quota, exchange rates, and political and economic stability are confounding factors that are

Table 1: Summary Statistics of Continuous and Discrete Variables

The variables *media* and $\tilde{media}$ show how positive or negative news articles are about the acquirer. A value close to one indicates that most articles are positive about the acquirer. A value close to zero indicates that most articles are negative. Sections 5.1 and 7 detail the construction of *media* and $\tilde{media}$, respectively. *aCash* denotes the acquirer's cash and the temporary investment vehicles for cash, including commercial paper and short-term government securities, as of the date of the most current financial information prior to the announcement of the transaction (mil. $). *aBookToMarket* is the book-to-market ratio of the acquirer. *aReturn* is the acquirer's share price return between the announcement date and the date four weeks prior to announcement. *days* is the number of days between announcement date and effective date in case of takeover success or date withdrawn in case of takeover failure. The columns labeled P_k show the kth percentile. The column labeled *Std. Dev.* shows the standard deviation.

Variable	Min	P_{25}	Mean	Median	P_{75}	Max	Std. Dev.
media	0.06	0.92	0.92	0.97	0.99	1.00	0.14
$\tilde{\text{media}}$	0.00	1.00	0.97	1.00	1.00	1.00	0.17
aCash (mil. $)	0.67	125.02	1418.03	420.73	1255.78	21971.00	2683.20
aBookToMarket	0.00	0.13	0.35	0.27	0.46	2.64	0.34
aReturn	-0.67	-0.08	-0.01	-0.01	0.06	0.61	0.17
days	1.00	74.00	129.43	101.00	161.00	764.00	90.62

nontrivial to control for when focusing on the media. Challenged deals are excluded because media dynamics fundamentally differs in bidding wars. Finally, only M&A transactions for majority/remaining interest and tender offers are included. The sample consists of 348 takeover attempts. 320 takeovers are successful (the variable *status* has value *completed*) and 28 takeover attempts fail (*status* is *withdrawn*).

To capture the effect of merger waves and macroeconomic factors, this paper considers the variable *prevTakeovers* consisting for each takeover attempt of the number of successful takeovers in the previous 100 days. Deal and firm characteristics are from SDC. They encompass the usual variables in the standard literature on mergers and acquisitions. News articles are from Dow Jones Factiva. For each takeover attempt, I include those articles that contain the names of the acquirer and the target within the first one hundred words. I collect articles that appear between the announcement date and one day prior to the effective date or the date withdrawn. Omitting the effective date or the date withdrawn ensures that any variable derived from news articles is exogenous with respect to the takeover outcome. A *text document collection* is the set of all included news articles. The resulting text document collection contains 82,830 news articles. Tables 1 and 2 show summary statistics and definitions of the most important variables.[9]

[9] Tables 1 and 2 show the subset of variables that are statistically significant in Section 6.

Table 2: Summary Statistics of Nominal Variables

status indicates whether the target or acquirer has terminated its agreement, letter of intent, or plans for the acquisition or merger or whether the transaction has closed. *stockSwap* indicates whether or not the acquiring company exchanges equity in itself for equity in the target. The acquirer must be acquiring at least 50% of the target's equity or be acquiring the remaining interest up to 100% of the target's equity, and at least 50% of the consideration offered must be in the form of equity. *unsolicited* indicates whether or not an acquiring company makes an offer for another company without prior negotiations.

Variable	Levels	Observations	%
status	withdrawn	28	8.0
	completed	320	92.0
stockSwap	no	134	38.5
	yes	214	61.5
unsolicited	no	322	92.5
	yes	26	7.5

5 Methodology

5.1 Constructing the Media Variable

This section shows how this paper quantifies text-based information. The first step consists of preprocessing the text document collection. Preprocessing entails conversion to lowercase, removal of stop words (examples include *as* and *the*), and stemming. Stemming is the process of erasing word suffixes to retrieve their radicals. It reduces complexity without significant loss of information.

The second step consist of creating a so-called *term-document matrix* from the preprocessed text document collection. This process is best illustrated by an example. Consider a text document collection[10] consisting of the two pangrams

1. the quick brown fox jumps over the lazy dog.

2. the jay, pig, fox, zebra, and my wolves quack!

Counting the occurrence of each word (*term*) in each text document yields the term-document matrix

	and	brown	dog	fox	jay	jumps	lazy	my	over	pig	quack	quick	the	wolves	zebra
1	0	1	1	1	0	1	1	0	1	0	0	1	2	0	0
2	1	0	0	1	1	0	0	1	0	1	1	0	1	1	1

[10]For clarity of exposition, this text document collection is only preprocessed by lowercase conversion.

Rows pertain to news articles and columns pertain to terms. The matrix element $f_{i,j}$ in row i and column j shows the frequency of term j in news article i.

With 116,901 columns, the resulting term-document matrix is very large. Many nonsensical terms such as *aaacplus* or *zzilr* are included. These terms are artifacts and are included because they occur at least once in at least one text document. Removing sparse terms eliminates these artifacts. It reduces the matrix dramatically without losing significant relations inherent in the matrix. Removing columns that have at least a 99 percentage of terms occurring zero times in a document yields a term-document matrix with 2,012 columns.

The third step consists of estimating a model that relates the news articles represented by the term-document matrix and their content. This paper uses the standard *naïve Bayes* model. The motivation for using this model is based on three of its properties. First, naïve Bayes works very well when tested on actual data sets (Cowell et al. (2007)). Second, it is computationally simple. This property is important because of the large dimensional nature of text-based information. Third, out-of-sample prediction of the variable of interest works very well (Antweiler and Frank (2004)).

Let C_i denote a random variable that has value one if news article i contains positive information about the acquirer and value zero otherwise. Let $F_{i,j}$ denote random variables corresponding to the realizations $f_{i,j}$, $j = 1, \ldots, 2012$. That is, $F_{i,j}$ pertains to the frequency of term j in the news article i. Then Bayes' rule implies[11]

$$\begin{aligned}
&P(C_i = c | F_{i,1} = x_1, \ldots, F_{i,2012} = x_{2012}) \\
&\propto P(C_i = c) P(F_{i,1} = x_1, \ldots, F_{i,2012} = x_{2012} | C_i = c) \\
&= P(C_i = c) \prod_{l=1}^{2012} P(F_{i,l} = x_l | C_i = c, F_{i,1} = x_1, \ldots, F_{i,l-1} = x_{l-1})
\end{aligned}$$

for all $c \in \{0, 1\}$ and all $x_1, \ldots, x_{2012} \in \mathbb{N}_0$, where $\mathbb{N}_0$ is the support of $F_{i,l}$. The key assumption of naïve Bayes is that the $F_{i,j}$'s are conditionally independent[12] given C_i. Thus,

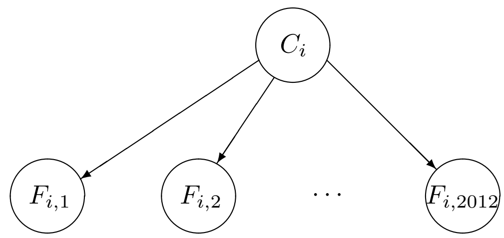
$$\begin{aligned}
P(C_i = c | F_{i,1} = x_1, \ldots, F_{i,2012} = x_{2012}) &\propto P(C_i = c) \prod_{l=1}^{2012} P(F_{i,l} = x_l | C_i = c) \\
= P(C_i = c) \frac{(\sum_{l=1}^{2012} x_l)!}{\prod_{l=1}^{2012} x_l!} &\prod_{l=1}^{2012} P(\text{obtain term } l \,|\, C_i = c)^{x_l}. \qquad (4)
\end{aligned}$$

The last equality follows because the term-document matrix implies a multinomial distribution. The probability $P(\text{obtain term } l \,|\, C_i = c)$ is the probability of obtaining term l when sampling from documents in category c. Figure 2 shows a path diagram illustrating the conditional dependencies implied by naïve Bayes.

[11] If $l = 1$, then $P(F_{i,l} = x_l | C_i = c, F_{i,1} = x_1, \ldots, F_{i,l-1} = x_{l-1}) := P(F_{i,1} = x_1 | C_i = c)$

[12] This assumption motivates the use of the term "naïve" in the naming of this model. Despite making this simplifying assumption, naïve Bayes often performs better with actual data sets than more sophisticated models.

Figure 2: Path Diagram Showing Conditional Dependencies of Naïve Bayes



Estimate model (4) with maximum likelihood. Estimating naïve Bayes requires observable realizations of C_i. To generate these realizations, draw a random sample of size 339 from the text document collection and classify the news articles manually.[13] The following is an example of an article classified as $C_i = 1$, that is, an article containing positive information about the acquirer:

> Standard & Poor's Ratings Services assigned its 'A' senior unsecured debt rating to UnitedHealth Group Inc.'s $250 million, 3.75% notes maturing Feb 10, 2009, and $250 million, 4.75% notes maturing Feb. 10, 2014. The proceeds will be used to finance the cash portion of the acquisition of Mid Atlantic Medical Services Inc.
>
> The rating reflects UNH's extremely strong consolidated earnings profile as well as its very strong financial flexibility and business position. The business position is enhanced by good product and market diversification and the company's unique competencies in specialized services. [...]

An example of an article classified as containing negative information about the acquirer ($C_i = 0$) follows:

> The modern personal computer is a technological marvel, but not such a good business proposition. Your desktop contains more computing power than mission control had for the Apollo moon landings, but the makers of the wizardry are laying off thousands of people and losing money. Hewlett-Packard Co. said this week that it would buy Compaq Computer Corp. in a bid to build a profitable business around the personal computer. Wall Street turned thumbs down immediately, knocking 18 percent off the value of H-P shares and 10 percent off Compaq shares. [...]

After model estimation using the random sample of size 339, calculate posterior probabilities

$$p_i := P(C_i = 1 | F_{i,1} = f_{i,1}, \ldots, F_{i,2012} = f_{i,2012})$$

[13] The number 339 results from drawing a sample of size 400 and discarding articles that do not directly pertain to the takeover in consideration. Most frequently, excluded articles are general market commentaries.

for all articles $i = 1, \dots, 82830$. The probability p_i is calculated out-of-sample except for the small estimation sample.

The fourth and final step consists of aggregating text-based information pertaining to a given takeover attempt. Several news articles are written for each takeover attempt. Interest lies in a variable summarizing the financial media's overall content for a given takeover. Letting $S_a \subseteq \{1, \dots, 82830\}$ denote the set of articles[14] pertaining to takeover attempt a, the variable *media* obtains as

$$\text{media}_a := \frac{1}{|S_a|} \sum_{i \in S_a} p_i, \qquad a = 1, \dots, 348, \tag{5}$$

where $|\cdot|$ denotes the cardinality of a set. The variable *media* is the average over all posterior probabilities pertaining to a given takeover attempt. If the majority of news articles carries positive information about the acquirer, *media* is close to one. If most articles are negative, *media* is close to zero. The exclusion of the effective date or the date withdrawn ensures exogeneity of *media* with respect to the takeover outcome *status*.

5.2 Binary Outcome Model

Takeover outcome is a binary variable. To operationalize the empirical prediction from Section 3, consider the general binary outcome model

$$P(\text{status}_a = \text{completed} \,|\, (\text{media}_a, \boldsymbol{x}_a)) = F((1, \text{media}_a, \boldsymbol{x}_a) \cdot \gamma). \tag{6}$$

The function F is the inverse of the probit or logit link,[15] $\boldsymbol{x}_a$ is a $1 \times K$ vector of control variables, and γ is the $(K+2) \times 1$ parameter vector to be estimated. If the empirical prediction of Section 3 holds, γ_2 is significant and positive and the sample average of the marginal effect

$$\frac{\partial P(\text{status}_a = \text{completed} \,|\, (\text{media}_a, \boldsymbol{x}_a))}{\partial \text{media}_a} = F'((1, \text{media}_a, \boldsymbol{x}_a) \cdot \gamma)\gamma_2$$

is positive.

6 Results

Figure 3 shows the spinogram relating *status* and *media*. It confirms graphically that the empirical prediction of Section 3 holds. Positive news articles about the acquirer tend to precede successful takeovers. Vice versa, negative articles about the acquirer precede failed takeovers.

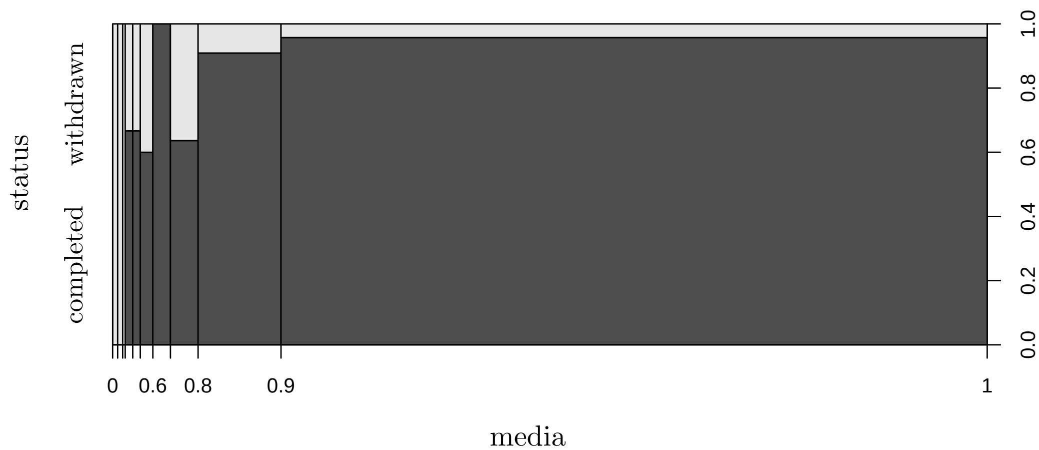
To ensure that this result is not caused by confounding variables, estimate model (6) with maximum likelihood. The control variables in $\boldsymbol{x}_a$ include deal characteristics, acquirer and target

[14] By construction of the text document collection, it holds that $S_{a_1} \cap S_{a_2} = \emptyset$ for all $a_1 \neq a_2$. This statement does not imply that no article pertains to more than one takeover. Instead it refers to the numbering of articles.

[15] $F(x) = \int_{-\infty}^{x} \phi(z)\mathrm{d}z$ or $F(x) = e^x/(1+e^x)$, respectively, where ϕ is the probability density function of the standard normal distribution.

Figure 3: Positive News Articles About Acquirer Entail Takeover Success

This figure shows that positive news articles about the acquirer (*media* is large) precede successful takeovers. Vice versa, negative articles about the acquirer (*media* is small) precede failed takeovers. The figure groups *media* in intervals and produces a spine plot for the resulting proportions of *status* within the *media* groups. Dark areas correspond to successful takeovers (*status* is *completed*). Light areas correspond to failed takeovers (*status* is *withdrawn*). The horizontal axis is distorted. The width of each *media* group is proportional to the corresponding number of observations. The large dark area on the right-hand side follows from the large amount of successful takeovers coinciding with $media \in [0.9, 1]$.



firm characteristics,[16] and *prevTakeovers*. The variable *prevTakeovers* is included to account for the possibility that a takeover might be more likely to succeed during merger waves.

Table 3 shows the estimation results after the elimination of variables with insignificant coefficients. Three key findings obtain. First, both probit and logit specifications of the link function yield a highly significant and positive *media* coefficient. This result confirms this paper's empirical prediction from Section 3. The average of the sample marginal effects are of similar magnitude in both model specifications.[17] An increase in one unit of *media* yields an increase in approximately 0.25 units of the probability that the takeover succeeds. Second, the addition of the single variable *media* to the model specification drastically increases goodness of fit. The inclusion of *media* raises the pseudo-R^2 by 42% and 38% compared to probit and logit models without *media*, respectively. Third, only firm characteristics pertaining to the acquirer are significant while target firm characteristics are insignificant. This result confirms the game-theoretic model's assumption that mainly uncertainty about the acquirer matters for shareholders' decision on whether or not to approve the takeover. It shows that the informational asymmetry facing target shareholders mainly pertains to the acquirer and not to target management.

The remaining coefficients are as expected. If the acquirer has ample cash reserves, if its market value is high relative to its book value, or if the acquirer's stock performs well prior to the takeover attempt, the takeover is likely to succeed. If the target intends to exchange equity in itself for equity in the target or if the takeover attempt is unsolicited, the takeover is likely to fail. Compared to unsuccessful takeover attempts, successful takeovers go along with a longer duration between announcement date and the date of resolution.

Figure 4 shows that *prevTakeovers* captures the cyclical nature of merger activity. Still its coefficient is insignificant in Table 3. There are two potential explanations for this phenomenon. The first explanation is that factors causing a merger wave influence the decision on whether or not to start a takeover attempt ex ante only. Once the takeover attempt has started, other factors become important. That is, the factors causing a takeover wave lose their importance ex post. This explanation is supported by the marginal increase in pseudo-R^2 when including *prevTakeovers* as an additional explanatory variable in model (6). Second, it is possible that the

[16]Deal characteristics include the number of days between announcement date and effective date or date withdrawn, a stock swap dummy, a dummy indicating the presence of anti-takeover devices, a tender offer dummy, a dummy indicating whether or not negotiations preceded an offer (*unsolicited* dummy), a proxy fight dummy, deal value, deal value to EBITDA, deal value to net sales, toehold, runup, markup, and premium (runup + markup). Firm characteristics include price to earnings, earnings per share, EBITDA to total assets, working capital to total assets, net income to net sales, price to sales, cash, cash to total assets, common equity, market value of equity, book to market, leverage, size, and share price return between announcement date and the date four weeks prior to announcement.

[17]The estimated coefficients differ in magnitude as a result of the different model specifications. This result is to be expected. Interpretation of coefficients outside the context of marginal effects is of limited value.

[18]The number of observations is slightly smaller than the number of takeovers in the sample because of missing values in some firm characteristics. Missing values only pertain to the sub-sample of successful takeovers. This sub-sample is large relative to its complement (Table 2). Only a minimal amount of information is therefore lost.

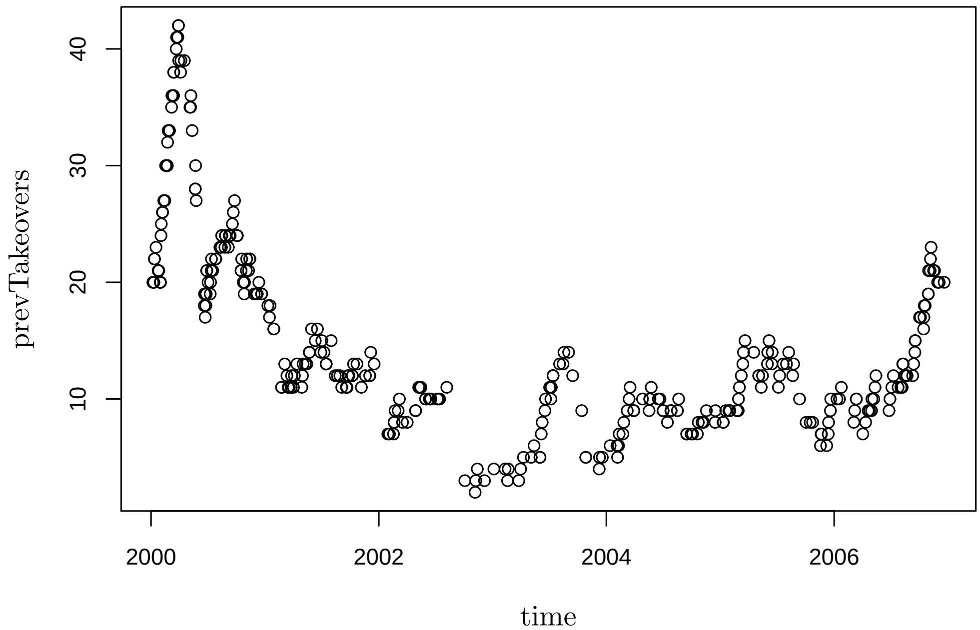
Table 3: Media Content Predicts Takeover Outcomes

This table shows the results of maximum likelihood estimation of model (6). The dependent variable equals one if the takeover *status* is *completed* and zero if it is *withdrawn*. Independent variables are described in Tables 1 and 2. The first row shows the specification of the link function. Columns labeled *Coeff.* show the estimated model coefficients, columns labeled *Marg. Effect* show the average of the sample marginal effects. "= *yes*" indicates the inclusion of a dummy variable with value equal to one if the nominal variable is equal to *yes* and zero otherwise. t statistics are in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively. The last two rows show McFadden's pseudo-R^2 and the number of observations.[18]

	Probit		Logit	
Variable	Coeff.	Marg. Effect	Coeff.	Marg. Effect
Intercept	-5.85***	-0.36	-10.65***	-0.26
	(-3.10)		(-2.98)	
media	4.69***	0.29	8.66***	0.21
	(4.99)		(4.88)	
log(aCash)	0.23**	0.01	0.46***	0.01
	(2.53)		(2.68)	
aBookToMarket	-1.45***	-0.09	-2.79***	-0.07
	(-3.34)		(-3.16)	
aReturn	3.43***	0.21	6.41***	0.16
	(3.30)		(3.16)	
stockSwap = yes	-1.64***	-0.10	-3.28***	-0.08
	(-2.80)		(-2.80)	
unsolicited = yes	-2.10***	-0.13	-4.11***	-0.10
	(-4.12)		(-4.04)	
log(days)	0.97***	0.06	1.76***	0.04
	(3.12)		(2.90)	
Pseudo-R^2	0.61		0.60	
Observations	332		332	

Figure 4: Merger Waves

This figure shows how the variable *prevTakeovers* changes over time. *prevTakeovers* shows for each takeover attempt the number of successful takeovers in the previous 100 days. The figure shows the peak of the merger wave in 2000, the subsequent decline, and the resumption of merger activity starting in 2003.



coefficient is insignificant because of the relatively small sample size.

7 Robustness Checks

Table 4 shows the results of robustness checks. The first robustness check in the columns labeled *Probit* shows an alternative construction of the *media* variable. As an alternative to (5), consider

$$\tilde{\text{media}}_a := \text{median}_{i \in S_a}\{p_i\}, \qquad a = 1, \ldots, 348.$$

Instead of aggregating text-based information pertaining to a given takeover using the mean, $\tilde{media}$ aggregates using the median. The result remains essentially unchanged.

The second robustness check in Table 4 shows an alternative specification of the inverse F of the link function. The motivation for this alternative is as follows. Table 2 shows that only 8% of takeovers fail (*status* is *withdrawn*). That is, one outcome is rare. In such cases the *complementary log-log model* is appropriate. Its inverse of the link function is the cumulative distribution function of the extreme value distribution.[19] It differs from the other models in

[19]That is, $F(x) = 1 - \exp(-\exp(x))$.

Table 4: Robustness Checks of Effect of Media on Takeover Outcome

The dependent variable equals one if the takeover *status* is *completed* and zero if it is *withdrawn*. Independent variables are described in Tables 1 and 2. The first row shows the specification of the link function. Columns labeled *Coeff.* show the estimated model coefficients, columns labeled *Marg. Effect* show the average of the sample marginal effects. "= *yes*" indicates the inclusion of a dummy variable with value equal to one if the nominal variable is equal to *yes* and zero otherwise. t statistics are in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively. The last two rows show McFadden's pseudo-R^2 and the number of observations.[18]

	Probit		Complementary Log-Log	
Variable	Coeff.	Marg. Effect	Coeff.	Marg. Effect
Intercept	-5.35***	-0.34	-6.50***	-0.58
	(-2.90)		(-3.46)	
media			5.04***	0.45
			(4.59)	
$\tilde{\text{media}}$	3.66***	0.23		
	(4.63)			
log(aCash)	0.26***	0.02	0.21**	0.02
	(2.77)		(2.36)	
aBookToMarket	-1.51***	-0.10	-1.44***	-0.13
	(-3.39)		(-3.36)	
aReturn	3.52***	0.22	3.56***	0.32
	(3.32)		(3.25)	
stockSwap = yes	-1.60***	-0.10	-1.90***	-0.17
	(-2.83)		(-2.75)	
unsolicited = yes	-2.61***	-0.17	-2.18***	-0.20
	(-5.06)		(-3.72)	
log(days)	0.99***	0.06	1.01***	0.09
	(3.13)		(3.39)	
Pseudo-R^2	0.60		0.62	
Observations	332		332	

being asymmetric around zero.

As with previous specifications, the coefficient of *media* is significant and positive. The complementary log-log model confirms the empirical prediction from Section 3. The main difference to Table 3 is that the averages of the sample marginal effects are larger in absolute value.

8 Conclusion

This paper shows both theoretically and empirically that the financial media plays a key role in mergers and acquisitions. The role of the financial media is to give the acquirer the opportunity to run a costly media campaign. Reading news articles, target shareholders obtain a noisy signal about the acquirer. When shareholder uncertainty is large and when the costs of running a media campaign are intermediate, a separating equilibrium obtains. The good acquirer type runs a media campaign with probability one to distinguish itself from the bad type. The bad acquirer type abstains from a media campaign because it is too costly to mimic the high type. Since the good and the bad type play different media strategies, the media signal is informative. If shareholders observe a positive media signal, they tender because they believe the acquirer generates value. The takeover succeeds. Vice versa, the takeover fails after a bad signal because target shareholders believe the acquirer destroys value.

Appendix

Proof of Lemma 1. Standing at the beginning of the game, the acquirer's expected payoff is $\pi := \beta\pi^H + (1-\beta)\pi^L$, where $\pi^t := P(S=0|T=t)\pi^{t,0} + P(S=1|T=t)\pi^{t,1}$ and $\pi^{t,s} := -c\xi_t + z(\mathbb{1}_{p\le b\le p^s} + \mathbb{1}_{p\vee p^s<b}) + (p_t - b)(k\mathbb{1}_{p\le b\le p^s} + n\mathbb{1}_{p\vee p^s<b})$. Observe that $P(S=s|\,T=t) = P(T=t|\,S=s)P(S=s)/P(T=t)$. With β^s and ζ from the proof[20] of Lemma 2, it follows that $P(S=1|T=H) = \zeta\beta^1/\beta = \xi_H\delta + (1-\xi_H)(1-\delta)$, $P(S=0|T=H) = (1-\zeta)\beta^0/\beta = \xi_H(1-\delta) + (1-\xi_H)\delta$, $P(S=1|T=L) = \zeta(1-\beta^1)/(1-\beta) = \xi_L\delta + (1-\xi_L)(1-\delta)$, $P(S=0|T=L) = (1-\zeta)(1-\beta^0)/(1-\beta) = \xi_L(1-\delta) + (1-\xi_L)\delta$. Thus, π as a function of b is piecewise linear and piecewise continuous. The only discontinuity occurs at $b=p$ since the potential discontinuities originating from $\pi^{H,s}$ and $\pi^{L,s}$ at $b=p^s>p$ cancel out. $\partial\pi/\partial b = 0$ on $[0,p)$ and $\pi(b_2) < \pi(b_1)$ for all $b_1, b_2 \in [p,\infty)$ with $b_1 < b_2$ (the function $\pi(b)$ is not necessarily differentiable on the whole interval (p,∞) because of piecewise linearity). Thus, $b^* = p$ maximizes π. $\square$

[20] This argument does not involve circular reasoning since Lemma 2 does not depend on results from Lemma 1.

Proof of Lemma 2. For $s \in \{0,1\}$ consider the posterior

$$\begin{aligned}
\beta^s &= P(T = H|\, S = s) \\
&= P(T = H|\, M = 1, S = s)P(M = 1|\, S = s) + P(T = H|\, M = 0, S = s)P(M = 0|\, S = s) \\
&= P(T = H|\, M = 1, S = s)\frac{P(M = 1, S = s)}{P(S = s)} + P(T = H|M = 0, S = s)\frac{P(M = 0, S = s)}{P(S = s)} \\
&= \frac{1}{P(S = s)}\big[P(T = H, M = 1, S = s) + P(T = H, M = 0, S = s)\big] \\
&= \frac{1}{P(S = s)}\big[P(T = H, S = s|\, M = 1)P(M = 1) + P(T = H, S = s|\, M = 0)P(M = 0)\big].
\end{aligned}$$

The distribution of S is conditional only on M and the distribution of M is conditional only on T. In particular,the distribution of S does not directly depend on T. Thus, T and S are independent conditional on M. That is, $P(T = t, S = s|\, M = m) = P(T = t|\, M = m)P(S = s|\, M = m)$. With $\zeta := P(S = 1) = P(S = 1|\, M = 1)P(M = 1) + P(S = 1|\, M = 0)P(M = 0) = \delta\mu + (1-\delta)(1-\mu)$, where $\mu := P(M = 1) = P(M = 1|\, T = H)P(T = H) + P(M = 1|\, T = L)P(T = L) = \xi_H\beta + \xi_L(1-\beta)$, it follows that

$$\begin{aligned}
\beta^1 &= \frac{1}{\zeta}\big[P(S = 1|\, M = 1)P(T = H|\, M = 1)P(M = 1) \\
&\quad + P(S = 1|\, M = 0)P(T = H|\, M = 0)P(M = 0)\big] \\
&= \frac{1}{\zeta}\big[\delta\frac{P(M = 1|\, T = H)P(T = H)}{P(M = 1)}P(M = 1) \\
&\quad + (1-\delta)\frac{P(M = 0|\, T = H)P(T = H)}{P(M = 0)}P(M = 0)\big] \\
&= \frac{1}{\zeta}\big[\delta\xi_H\beta + (1-\delta)(1-\xi_H)\beta\big] = \beta\cdot\frac{\delta\xi_H + (1-\delta)(1-\xi_H)}{\zeta}.
\end{aligned}$$

The expression for β^0 follows from an argument analogous to the derivation of β^1. □

Proof of Lemma 3. The first claim follows from $p_L < p < p_H$ and $\delta \in (1/2, 1)$. The second claim follows from $-(p - p_L) < 0 < (p_H - p)$. □

Proof of Theorem. Lemma 1 implies that $\pi^{t,s} = -c\xi_t + [z + k(p_t - p)]\mathbb{1}_{p \le p^s}$ (see the proof of Lemma 1 for the definitions of $\pi^{t,s}$ and π^t). The key to this proof is the acquirer's expected profit $\pi^t = \pi^t(\xi_t)$ at the point in time when it decides on the media strategy. The high type with objective function π^H plays against the low type with objective function π^L. The goal is to find a perfect Bayesian equilibrium in this signaling game. That is, find a fixed point (ξ_H^*, ξ_L^*) such that $\xi_t^* \in \operatorname{argmax}_{\xi_t} \pi^t(\xi_t)$, where p^s in the function $\pi^t(\xi_t)$ is evaluated at (ξ_H^*, ξ_L^*).

This proof's focus is on the separating equilibrium. The pooling equilibrium follows from an analogous argument. Resolve shareholder indifference by letting shareholders tender if $p = p^s$. It holds that $\beta \ge \underline{\beta}$ if and only if $p \le p^1|_{(\xi_H,\xi_L)=(1,0)}$. It also holds that $\beta \ge \bar{\beta}$ if and only if $p \le p^0|_{(\xi_H,\xi_L)=(1,0)}$. Let $\beta \in [\underline{\beta}, \bar{\beta})$. Then $p^0|_{(\xi_H,\xi_L)=(1,0)} < p \le p^1|_{(\xi_H,\xi_L)=(1,0)}$ holds. Thus, shareholders tender after observing $s = 1$ and they do not tender after observing $s = 0$ if the acquirer plays $(\xi_H, \xi_L) = (1, 0)$. With $p^s = p^s|_{(\xi_H,\xi_L)=(1,0)}$, it follows that $\pi^L(1) \le \pi^L(0)$ if and only if $c \ge \underline{c}$, and that $\pi^H(0) \le \pi^H(1)$ if and only if $c \le \bar{c}$. Thus, the desired result follows. □

References

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(3):1259–1294.

Bagnoli, M. and Lipman, B. L. (1988). Successful Takeovers without Exclusion. *Review of Financial Studies*, 1(1):89–110.

Cowell, R. G., Dawid, A. P., Lauritzen, S. L., and Spiegelhalter, D. J. (2007). *Probabilistic Networks and Expert Systems: Exact Computational Methods for Bayesian Networks*. Information Science and Statistics. Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA.

Diamond, D. W. and Dybvig, P. H. (1983). Bank Runs, Deposit Insurance, and Liquidity. *Journal of Political Economy*, 91(3):401–419.

Feinerer, I., Hornik, K., and Meyer, D. (2008). Text Mining Infrastructure in R. *Journal of Statistical Software*, 25(5).

Grossman, S. J. and Hart, O. D. (1980). Takeover Bids, The Free-Rider Problem, and the Theory of the Corporation. *Bell Journal of Economics*, 11(1):42–64.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I. H. (2009). The WEKA Data Mining Software: An Update. *SIGKDD Explorations*, 11(1).

Harris, M. and Raviv, A. (1993). Differences of Opinion Make a Horse Race. *Review of Financial Studies*, 6(3):473–506.

Hornik, K., Buchta, C., and Zeileis, A. (2009). Open-Source Machine Learning: R Meets Weka. *Computational Statistics*, 24(2):225–232.

Ohl, C. M., Pincus, J. D., Rimmer, T., and Harrison, D. (1995). Agenda Building Role of News Releases in Corporate Takeovers. *Public Relations Review*, 21(2):89–101.

Tetlock, P. C. (2007). Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *Journal of Finance*, 62(3):1139–1168.

Tetlock, P. C., Saar-Tsechansky, M., and Macskassy, S. (2008). More Than Words: Quantifying Language to Measure Firms' Fundamentals. *Journal of Finance*, 63(3):1437–1467.

Veldkamp, L. L. (2006). Media Frenzies in Markets for Financial Information. *American Economic Review*, 96(3):577–601.