Essays on the Media's Production and Dissemination Role:

Evidence from financial and non-financial disclosure

Dissertation

zur Erlangung des Doktorgrades

der Wirtschaftswissenschaftlichen Fakultät

der Georg-August-Universität Göttingen

vorgelegt von

Ann-Kristin Großkopf

geboren in Göttingen

Göttingen, 2020

Betreuungsausschuss:

Erstbetreuer:	Prof. Dr. Jörg-Markus Hitz	(Erstprüfer)
Weitere Betreuer:	Prof. Dr. Jan Muntermann	(Zweitprüfer)
	Prof. Dr. Maik Hammerschmidt	(Drittprüfer)

Tag der mündlichen Prüfung: 22.04.2020

Table of Contents

List of Tables

List of Figures

Page

1 Introduction

1.1 Context of the thesis

This thesis examines the information production and dissemination role of the media regarding financial and non-financial disclosure. In three empirical studies, the thesis address two motivating questions: (1) How does the financial press disseminate and produce information? And (2) how do stakeholders react to non-financial information dissemination through the press?

The first and the second study of the thesis (chapters three and four) address the first motivating question. While the first study focuses on the supply of quantitative information in press articles, the second study exploits the textual features of these articles, such as the topics or the sentiment. Both studies contribute to the literature related to the financial press as an intermediary of financial information. To that end, the thesis addresses and extends the debate about the value of financial press articles for investors and provides detailed evidence about the information production in those articles.

The third study (chapter five) shifts the focus with regard to the information dissemination from the perspective of investors and their need for financial information to another stakeholder group, namely employees and their information demand. For stakeholders, non-financial information is of great importance, as stakeholders such as customers or employees are at the core of non-financial activities (Christensen et al. (2019)). To provide answers related to the second motivating question, in chapter five, the study examines the short-term and long-term implications of increased stakeholder attention in response to non-financial information in the press. The evidence of the third study extends the scarce literature about the link between stakeholder theory and non-financial disclosure.

The following section presents the literature stream related to the financial press as an intermediary of financial information and the stream about the link between stakeholder theory and non-financial disclosure to define the terminology of each literature stream. Additionally, the section points out important gaps in the literature that the three empirical studies of this thesis aim to fill by answering the motivating questions.

Financial press as an intermediary of financial information

Beaver (1998, p. 10) defines information intermediaries as „an industry whose factors of production include financial information and other types of data and whose product is analysis and interpretation". Investors can choose from a vast array of information intermediaries in capital markets such as analysts, proxy advisors, or the financial press, which offer different information related products (Healy and Palepu (2001)). Though just one among many information intermediaries, the financial press traditionally has the broadest and most diversified audience of all common information intermediaries. Prior literature mainly discriminates two sources of value creation for press articles: information dissemination and information production.

Information dissemination

The financial press collects and rebroadcasts information about accounting events such as the publication of quarterly financial statements and makes this information available to their readers through the articles. This information can take various forms, such as earnings figures or quotes from firm executives. This task potentially saves transaction costs of information gathering on behalf of investors. Also, it creates common knowledge of the information among market participants (Dyck and Zingales (2003)). Part of the information gathering is the reduction of all information available to focus on key items such as performance measures, which also means that journalists only partly disseminate public information. The dissemination occurs via various news outlets either in the traditional print version or via the online version of the newspaper.

There is ample evidence supporting the dissemination role of the financial press. For example, Peress (2014) examines the impact of press coverage using the exogenous reduction in coverage due to a sample of national newspaper strikes in several countries. His results show a significant decline in trading volume and volatility on strike days and strongly support the information dissemination role of the press. This evidence complements the results of Engelberg and Parsons (2011) for local trading markets and is consistent with media contributing to the speed of information diffusion into prices. Bushee et al. (2010) examine the financial press coverage around earnings announcements and find that higher press coverage is associated with lower bid-ask spreads and increased market depth. This result indicates that press coverage reduces information asymmetry and thus contributes to market efficiency. Information dissemination by the financial press is particularly relevant for firms with a low information environment. To that end, Dyck and Zingales (2003) show that the market price impact of press articles is stronger for companies with low analyst following.

Recent evidence by Drake et al. (2014) documents that increased press coverage of annual earnings announcements mitigates cash flow mispricing, but not accrual mispricing. Detailed analyses reveal that this result only holds for newsflashes, i.e., pure rebroadcasting news excerpts but not for full articles, which might provide incremental journalist content. Blankespoor et al. (2018) examine computer-generated articles and thus eliminate any information production from the articles. The authors show that this kind of coverage increases market liquidity, while the speed of the price discovery seems rather unaffected. They attribute this effect to retail investors because they become aware of investment opportunities due to the “robo-coverage”. Consistent with this, Kothari et al. (2009) argue that the financial press is a more reliable source for (retail) investors compared to the firms or analysts. Financial press publications generally attract investor attention, even though the information provided may not be new or specifically value relevant. To that end, Madsen and Niessner (2019) show an increase in trading volume due to a higher level of investor attention following print advertisements in newspapers.

Information production

Complementary to the dissemination role, prior literature identifies an information production role of the financial press, which broadly encompasses the content that articles provide incremental to rebroadcasting information. There are two main sets of information with which the financial press can produce information. First, journalists can collect, aggregate, and combine information taken from various public sources. For example, the press may rely on the expertise of other information intermediaries like analysts (Rees et al. (2015); Guest and Jaewoo (2019)) or may discover the information in public court actions (Miller (2006)). This way, the press extends mere firm information and makes valuable but less broadly known information available to a larger audience. Second, journalists may discover new information through private interactions with company representatives or other experts in the respective field.

Additionally, the professional skills of journalists enable them to present the information consistent with the informational needs of their readers and to collect key items, which enhances the understanding of the (retail) investors. To create valuable content for readers, journalists need to feature items in their articles that are useful for the decision-making of investors with regard to the information content and materiality to avoid information overload.

By documenting abnormal returns upon the publication of press articles with original content, event studies provide evidence that the press assumes an information production function, e.g., with respect to identifying low-quality accounting (Foster (1979)), or fraud cases (Miller (2006)). Miller (2006) further introduces a 'watchdog function', which relates to the monitoring behavior of the press. This monitoring function helps investors to hold firms accountable, e.g., for excessive executive pay, corporate governance violations or insider trading (Core et al. (2008); Dyck et al. (2008); Bednar (2012); Dai et al. (2015)).

Tetlock et al. (2008) examine the sentiment of news articles and find that the fraction of negative words in an article is associated with lower future earnings on a firm level. Thus, the information production is not limited to the facts presented in the articles, but also includes the choice of words and the resulting sentiment as another article attribute. Building on these results, Guest (2018) finds that Wall Street Journal (WSJ) coverage of earnings announcements is associated with higher trading volume and more efficient price responses, especially when the article contains more original content. The author measures original content as the similarity between the earnings press release and the WSJ article.

While the studies mentioned above are aligned with the information dissemination and production role und thus support the role of the financial press as an information intermediary, prior literature also provides a more critical view of the financial press. Tetlock (2011) indicates that individual investors overreact to articles when they contain stale news (i.e., similar content (phrasing) as previous articles). The results of Ahern and Sosyura (2015) and Ahern and Sosyura (2014) further question the view of the financial press as an intermediary. The authors show that firms can influence their own press coverage by publishing more press releases. Furthermore, the press wants to appeal to its readership and thus publishes articles that might be less accurate and based on rumors. Cahan et al. (2015) show that firms manage their public image in the media through CSR activities. These results are in line with the argument of Jensen (1979) that the press mainly exists to entertain its readers. Gurun and Butler (2012) argue that news outlets might be subject to a conflict of interest when they report about a local firm or one with high (local) advertising expenditures resulting in biased reporting.

Overall, it remains an open empirical question how financial press articles create value for their audience, i.e., if the information dissemination or the information production role is economically more meaningful or if the value is indeed limited to the entertainment of readers. With regard to information production, prior literature also lacks evidence about how the press produces original content or what kind of editorial content is supplied.

Stakeholder theory and non-financial disclosure dissemination

Corporate social responsibility disclosure

Based on Christensen et al. (2019), this thesis defines corporate social responsibility (CSR) as "corporate activities and policies that assess, manage, and govern a firm's responsibilities for and its impacts on society and the environment". Consequently, CSR reporting refers to the disclosure of this kind of information. Prior literature shows that the dissemination of CSR information has value implications for the firms, which is comparable to financial information. For example, Flammer (2013) and Krüger (2015) examine short-term market reactions to positive and negative CSR events published in the press or the Kinder Lydenberg Domini (KLD) database and find that investors react negatively to the publication of negative non-financial information such as news about eco-harmful firm behavior. While Flammer (2013) finds that investors value news implying social responsibility (i.e., positive news), Krüger (2015) argues that investors only react favorably to the release of positive information when the events relate to resolving prior stakeholder relationship issues. Edmans (2011) documents positive long-term capital market effects and an increase in positive earnings surprises for firms signaling good CSR performance.

For CSR information published by the firms, e.g., CSR reports, the literature distinguishes between implications of mandatory and voluntary disclosure. With regard to voluntary disclosure, Dhaliwal et al. (2011) document high cost of capital as a motive to initiate the disclosure of stand-alone CSR reports for US firms. However, positive effects from issuing such a report, e.g., the reduction of cost of capital and higher analyst coverage, only materialize for firms whose CSR performance exceeds the industry median. For an international sample, Dhaliwal et al. (2012) find a positive association between the publication of a CSR report and the analyst forecast accuracy. Additionally, the results of Cahan et al. (2015) suggest a positive association between voluntary CSR disclosure and firm value, which is more pronounced in countries where non-financial information is less transparent, e.g., due to weak institutions. The results of Christensen (2016) imply that the voluntary publication of non-financial information is beneficial for firms, as it moderates the negative capital market effects if the reporting firm is charged with social or environmental misconduct in the following year creating a kind of “insurance function”.

In recent years, several jurisdictions implemented legislation requiring firms to publish CSR reports. For example, since 2017 the EU requires large firms to publish reports containing information about environmental matters, social and employee-related matters, respect for human rights, anti-corruption and bribery matters (Directive 2014/95/EU, recital 6). Grewal et al. (2019) examine market reactions for events leading to the EU directive and find an overall negative market reaction, which is muted for firms with higher CSR performance or disclosure before the new regulation. Fiechter et al. (2019) examine how firms react in anticipation of the EU disclosure mandate and find that EU firms increase their CSR activities compared to U.S. control firms. This increase is positively associated with the treatment intensity, i.e., the pre-regulation disclosure level and the resulting required increase in CSR disclosure. Chen et al. (2018) report a deterioration of performance and an increase in expenses following the implementation of mandatory CSR disclosure in China. However, the authors also report a reduction in environmental pollution in cities with a high proportion of treatment firms, which suggests that the regulation creates positive externalities.

In contrast to the EU and China, US regulation does not require firms to publish full-fledged CSR reports. However, several separate regulations demand the disclosure of specific CSR dimensions. For example, Christensen et al. (2017) examine the effects of implementing mandatory mine-safety disclosure in financial reports. Previously, the safety records were publicly available via the Mine Safety and Health Administration’s (MSHA) website. Following the implementation of the new disclosure rule, the authors find an increase in safety regulation compliance, which is in line with the documented decrease in injury rates. The results additionally show an increase in analyst and media attention, which suggests that the public awareness of the information depends on where it is published, i.e., how easily investors and other stakeholders can collect this non-financial information. Similarly, Hombach and Sellhorn (2019) examine the SEC requirement for oil and gas firms to disclose extraction payments. Using an event study methodology for the likelihood of this regulation, the authors find that investors react more negatively when the firms exhibit higher reputational risks from the disclosure requirement.

Stakeholder theory

Regulators like the EU stress that CSR disclosure should "meet the needs of investors and other stakeholders as well as the need to provide consumers with easy access to information on the impact of businesses on society" (Directive 2014/95/EU, recital 3). Therefore, stakeholders are an essential group to consider when examining CSR disclosure. Freeman (1984) defines a stakeholder as "any group or individual who is affected by or can affect the achievement of an organization's objectives". Important stakeholders are customers, employees, or shareholders. Prior literature further distinguishes between internal stakeholders with close ties to the firm such as investors or employees, and external stakeholders like customers (Wood (1995); Madsen and Rodgers (2015)). As non-investor stakeholders comprise rather diverse groups, it is difficult to examine all of them at once. Regarding CSR activities, employees are an essential stakeholder group of a firm. Therefore, this thesis focuses on employees as a specific group of stakeholders.

Firms can benefit from a strong commitment to their employees and their well-being, as the stakeholder hypothesis suggests that a strong bond between the employees and the firm strengthens the firms' reputation in the market and stimulates employee productivity. Additionally, a strong commitment reduces the risk of strikes or high fluctuation among highly qualified employees (Ben-Nasr and Ghouma (2018); Edmans (2011); Jones (1995)). As internal stakeholders, employees are subject to the internal and potentially long-term CSR activities, for example, how a firm tries to promote diversity among its employees. Additionally, employees are affected by the external non-financial performance, which has an impact on the entire society, for example, based on environmental issues.

While the link between stakeholders and non-financial information is an important, there are very few papers examining this link. One of these papers is Madsen and Rodgers (2015), who examine stakeholder attention to corporate disaster relief efforts. The authors measure attention using short-term and long-term newspaper coverage. The results show a positive association between the stakeholder attention and their measures for the legitimacy, urgency, and enactment of the relief programs, which ultimately results in financial benefits for the firms. However, the authors cannot provide evidence about the internal, long-term CSR activities of a firm and how the stakeholders react to it directly. Therefore, it remains an open empirical question if and how non-investor stakeholders use non-financial disclosure and what implication their use of it might have for the reporting firms.

1.2 Motivation and objective of the thesis

Based on the context of the thesis and the respective open empirical questions outlined above, this thesis examines the role of the financial press as an information intermediary and the link between stakeholders and non-financial information disseminated through the press. As outlined in Figure 1.1, the thesis examines two main motivating questions related to each literature stream and research gap.

Regarding the literature stream related to the financial press as an intermediary of accounting information, there is an ongoing debate about how journalists create value for their readers, especially when the articles feature information about earnings announcements. On the one hand, authors such as Drake et al. (2014) and Bushee et al. (2010) argue that the value of press articles is limited to information dissemination, while others such as Guest (2018) or Coyne et al. (2019) stress that the press also produces valuable information, i.e., editorial content.

If the press only assumed an information dissemination role, it would be arguably why the various news outlets spend resources on covering these earnings announcements instead of relying on automated articles as in Blankespoor et al. (2018). Additionally, mere information dissemination would not be in line with the journalists' perspective on their job. In the survey paper of Call et al. (2018), journalists state that they have incentives to provide exclusive content (e.g., from private conversations with managers or analysts) and high-quality articles, which supports that journalists want to produce supplemental information. Assuming that the financial press indeed produces valuable information, we know very little about the origin of such production because prior literature so far has only established fairly general proxies for information production. For example, Drake et al. (2014) and Coyne et al. (2019) measure information production based on a sample split into so-called newsflashes and full articles. While newsflashes do not produce any information, the information production in articles varies substantially.

Financial press articles feature various attributes, which can influence the way journalists produce information (Graf-Vlachy et al. (2019)). First, articles, especially those about earnings announcements, feature quantitative items. Second, press articles feature various textual characteristics such as the sentiment, various topics, and the readability. Guest (2018) uses such a textual feature of the articles and measures information production based on the similarity of earnings announcements and the related article. While this proxy is more related to the article itself than the one of Drake et al. (2014), it merely suggests if an article produces information without establishing how the information is produced. Therefore, it is an open empirical question how and what kind of news outlets produce information.

The aim of study 1 and study 2 of this thesis is to provide evidence regarding this question. As outlined in Figure 1.1, both studies have a similar motivating question. However, each paper examines a specific though related research question and uses a different article attribute and methodological approach to answer it. The first study features an explorative approach and uses the quantitative items of the articles based on a manual content analysis to answer the question about what kind of information various news outlets supply in their articles. An item-based reconciliation of the supplied information with firm disclosure and conference calls allows a direct distinction between information dissemination and production.

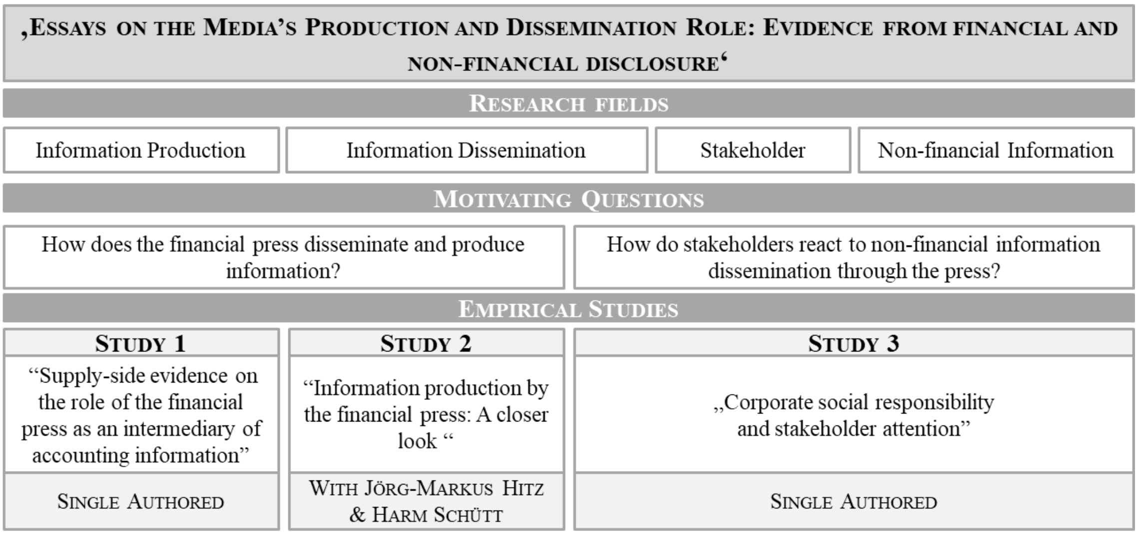


Figure 1.1: Context of the thesis

The second study examines what kind of, and how journalists produce information in their articles. To that end, the study exploits the textual features of the articles instead of quantitative items. The study uses a Bayesian topic modeling algorithm to provide evidence about the topics featured in the articles and to distinguish information production and dissemination. Additionally, the study captures the sentiment of each topic to get directional insights into the production of news outlets. Overall, both studies provide extensive evidence about the production in financial press articles based on their main attributes and therefore extend the prior literature about the potential value of financial press articles. Extending the first study, the second study additionally employs market tests to provide evidence about the potential value of information production for investors.

As outlined in Figure 1.1, the second motivating question aims at providing evidence about how stakeholders react to non-financial information dissemination through the press. Therefore, the information dissemination is the key link between the first two studies and the third one. Study 3 relies on the range and credibility of the press to disseminate the non-financial information to the target group. The third study focuses on non-investor stakeholders rather than investors, as this kind of information is of special interest to stakeholders. The dissemination through the press enhances the likelihood that stakeholders become aware of the CSR information because it is less likely that stakeholders use traditional disclosure channels such as SEC filings to collect information.

Stakeholders such as employees or customers are at the core of CSR, as they are often affected by the related firm actions (Christensen et al. (2019)). Though firms ought to consider stakeholders when disclosing non-financial information, e.g., based on the recent EU directive, it is an open empirical question if and how stakeholders actually use this information and how they react to its disclosure (Hombach and Sellhorn (2018)). This lack of evidence partly derives from the non-monetary interests of stakeholders. Additionally, the information demand of stakeholders changes over time, based on external influences such as the recent debate about climate change (Christensen et al. (2019)). However, it is essential to understand the information demand of stakeholders to draft effective legislation and provide useful information. Thus, the third study extends the still scarce literature about the stakeholders' use of and reaction to non-financial information and therefore provides valuable insights for users, regulators, and preparers.

1.3 Content of the thesis

The thesis comprises three studies. As outlined in Figure 1.2, the thesis proceeds as follows. The first study of this thesis investigates the supply of quantitative accounting information in financial press articles (Chapter 2). The second study examines whether and how useful the information that is provided to investors in financial press articles using a topic modeling approach (Chapter 3). The third and final study of this thesis provides evidence on short and long-term stakeholder reactions to non-financial disclosure (Chapter 4). The final chapter concludes (Chapter 5).

Chapter 2: Supply-side evidence on the role of the financial press as an intermediary of accounting information

The *first study* of this thesis examines the supply and content of financial articles following quarterly earnings announcements. Using a sample of S&P 500 firms between Q1-2013 and Q2-2016, the study provides explorative evidence about 1) the decision of an outlet to cover a firm, 2) the content of the articles, and 3) the extent to which journalists rely on sources other than the firms. The study analyses three different news outlets: The Wall Street Journal, as a specialized news outlet, and the New York Times and USA Today as general-interest outlets. Results from estimating a probit regression suggest that the coverage of quarterly earnings announcements is associated with the richness of the firms' information environment and the extent of public interest. Coverage decisions do not vary substantially between the three news outlets. A detailed content analysis of the financial press articles reveals that news outlets feature accounting items such as performance or revenue metrics prominently in their articles. Furthermore, the articles supply external information like analyst forecasts or stock market information. Non-accounting items only play a subordinate role in the sample articles.

The study further investigates the sources of the information supplied in the articles. The results suggest that items about firm performance and revenue emanate from firm disclosure and are therefore subject to information dissemination. Additionally, the analysis reveals that articles feature items from conference calls or other (private) sources, such as M&A activities and expenses. Therefore, the articles also produce new and potentially valuable information for the readers/investors. An analysis of the quotes featured in the articles further reveals that information production is not limited to "hard information" but extends to the verbal level. The amount of information production in financial press articles is again associated with firm fundamentals such as size or liquidity. Overall, the results of the study indicate that the financial press assumes both an information dissemination and an information production role.

1. Introduction

- Context of the thesis
- Motivation and objective of the thesis
- Content of the thesis

Pages 1-14

Main objective of the study	Main results of the study	Main rel. literature
Chapter 2: Supply-side evidence on the role of the financial press as an intermediary of accounting information		
Sheds light on the information production and dissemination role of the different outlets based on a content analysis of quantitative items	The financial press assumes an information dissemination and an information production role	Drake et al. (2014, TAR)
	Outlets decide to provide coverage and produce additional information based on the general information environment of a firm	Guest (2018, WP)

Pages 15-58

Main objective of the study	Main results of the study	Main rel. literature
Chapter 3: Information production by the financial press: A closer look		
Addresses how and what kind of information journalists produce in the narratives of in financial press articles using a topic modeling approach	Journalists use various article attributes such as content and sentiment to produce information	Guest (2018, WP)
	The produced content in the article has a stronger impact on the investors' response to earnings than the article sentiment	Huang et al. (2017, Management Science)

Pages 59-98

Main objective of the study	Main results of the study	Main rel. literature
Chapter 4: Corporate social responsibility and stakeholder attention		
Provides evidence on the long and short-term consequences of non-investor stakeholder attention to non-financial disclosure	Stakeholders pay attention to the disclosure of non-financial information but do not act upon it	Madsen et al. (2015, Strategic Management Journal)
	Firms increase their CSR activities in anticipation of increased stakeholder attention	Greening et al. (2000, Business & Society)

Pages 99-153

5. Conclusion

- Summary of major findings and implications
- Main limitations
- Avenues for future research

Pages 154-158

Figure 1.2: Content of the thesis

Chapter 3: Information production by the financial press: A closer look

The *second study* extends the insights about the role of the financial press as an information intermediary and focuses on how and what kind of information the press produces in its articles. In contrast to the first study, the second study examines article narratives and excludes quantitative items. Based on 6,540 WSJ articles about S&P 500 firms' earnings announcements between 2010 and 2016, the study uses a topic modeling approach to identify the original content of the press articles. Based on the topic model, the study identifies four sources of information production: 1) adding new topics, 2) omitting potentially irrelevant content from the earnings announcement, 3) emphasizing specific topics, and 4) the sentiment of the narratives.

Descriptive analyses show that financial press articles feature, on average, 1.5 additional topics and omit 5.5 topics from the earnings announcement. Journalists in the sample emphasize key performance topics such as earnings and sales and de-emphasize topics like non-GAAP earnings or firm segments. While the sentiment of press articles is generally lower compared to earnings announcements, the sentiment of topics added to the articles is, on average, higher than the sentiment of topics also featured in the earnings announcement.

Multivariate analyses reveal that both adding and omitting topics in financial press articles are associated with larger investor responses to the announced earnings. This positive association increases with the level of information production, i.e., with articles that feature more original content or omit more potentially irrelevant content. Interestingly, only the sentiment of topics featured in the earnings announcement and the corresponding press article has a positive association with the investor response, while the sentiment of additional topics does not. These results suggest that what journalists report in their articles is more important than how they phrase it. Overall, the results of the study support that the financial press creates additional content for its audience. Thus, the results suggest that the value of the articles exceeds the mere entertainment of the readers.

Chapter 4: Corporate social responsibility and stakeholder attention

The *third and final study* of this thesis investigates the long and short-term consequences of non-investor stakeholder attention to non-financial disclosure. It uses a sample of Fortune's "100 best companies to work for" list between 2005 and 2017 as a specific form of non-financial disclosure. This list ranks firms based on an employee and a culture audit survey and rates areas such as respect, fairness, or benefit programs. This setting offers three main advantages. First, the dissemination of the financial information occurs through the media as an information intermediary, which enhances the credibility of the information. Additionally, the disclosure through this channel makes it more likely that the stakeholders become aware of the information. Second, there are two key dates for the publication of the list data. The first date is the first disclosure of the Fortune information, e.g., through the Fortune web page, while the second event is the official issue date of the Fortune (print) issue. Thus, the information is disseminated via two different channels on two different dates, which facilitates the separation of stakeholder and investor effects. Third, the Fortune list contains information about the CSR performance of the firms regarding a specific non-investor stakeholder group, which makes it possible to isolate the reaction of specific stakeholders. This, in turn, simplifies the identification of the effects.

The study answers two different research questions. First, it examines whether stakeholders increase their attention to a firm after the release of externally validated non-financial information, as they need to be aware of the information to react to it. It measures the stakeholder attention using the abnormal Google search volume index based on firm name searches and finds significant increases in stakeholder attention following both the first disclosure of the list and its wider dissemination through the Fortune issue. A comparison with the investor reaction shows that only the first event contains valuable information content for investors.

The second research question examines how stakeholder attention may have long-term consequences for firms based on stakeholder reactions and their own actions. To that end, the paper examines if stakeholders react to the list information by applying for jobs at the list firms. It finds an increase in job applications, which is not statistically significant. Using a difference-in-difference analysis based on a matched sample of list and non-list firms, the paper shows that firms increase their CSR activities once they are featured on the list. Overall, the results of this study suggest that stakeholders pay attention to the publication of stakeholder related non-financial information but do not act upon it, whereas the firms invest in their CSR activities in anticipation of the stakeholder attention.

2 Supply-side evidence on the role of the financial press as an intermediary of accounting information

Ann-Kristin Großkopf[1]

Working Paper[2]

Abstract: I shed light on the economic role of the financial press as an intermediary of accounting information by providing explorative, descriptive evidence on the supply of newspaper coverage and content. My content analyses of three different news outlets offer detailed insights into the structure of information provided and hence helps to understand factors that shape demand for newspaper coverage. I find that the supply of financial press information is associated with the quality of the firm's information environment and that media attention focuses on firm size. My analyses of the sources of information contained in press articles also reveal that journalists systematically provide additional information beyond re-broadcasting items taken from firms' earnings announcements. I contribute to the literature by providing an explorative supply-side perspective on the role of the financial press as an information intermediary, which is consistent with the press assuming both an information dissemination and a production role.

JEL Classification: M40, M 41.

Keywords: Financial press, information intermediation, earnings announcement

Acknowledgments: For valuable comments, I am grateful to session participants at the Annual Meeting of the European Accounting Association in Valencia, Spain (May 2017), and my dissertation adviser Jörg-Markus Hitz.

[1] Georg-August Universität Göttingen, Faculty of Economic Sciences, Chair of Accounting and Auditing, Platz der Göttinger Sieben 3, 37073 Göttingen, Germany.

[2] First version of the working paper: November 2016. This version: December 2019.

2.1 Introduction

The financial press is an important information intermediary in capital markets, which presumably has the broadest audience of all intermediaries. Prior research shows that the dissemination of firm information through this channel increases trading volume and market liquidity (e.g., Blankespoor et al. (2018); Engelberg and Parsons (2011)). Additionally, the press has got the potential to shape the view of its readers and trigger significant market reactions. For example, Wirecard AG shares plunged up to 30% when the Financial Times published an article accusing the firm of "money laundering" and "falsification of accounts" in January 2019. Following these reports, Wirecard filed a lawsuit against the Financial Times, accusing the outlet of unethical reporting and market manipulation. While cases like this show how important it is that (financial) news outlets publish trustworthy and high-quality information, we know little about how the press creates value for its readers. Prior literature provides contrasting evidence attributing the value of financial press articles merely to information dissemination (e.g., Drake et al. (2014)), while others stress the information production (e.g., Guest (2018)). I contribute to this stream of literature by providing an explorative supply-side perspective based on a content analysis. My approach differs from prior research, as it allows me to divide all quantitative items into information dissemination, i.e., items based on firm disclosure and production, i.e., items from other (public) sources. Such an analysis promises to provide new insights into the demand for the services of the financial press and provides a new angle on how these services potentially create value based on the production and dissemination role.

Following Bushee et al. (2010), this paper defines an information intermediary as "an agent that provides information that is new and useful to other parties, either because it has not previously been publicly released or because it has not been widely disseminated" (Bushee et al. (2010), p. 1). These tasks potentially save transaction costs for investors. As outlined above, prior literature broadly distinguishes two roles of the financial press as an intermediary of accounting information: information dissemination and information production. By assuming an information dissemination role, the financial press helps to broadcast information released by firms, e.g., to the target audience such as investors, hence contributing to the proliferation of information and its recognition in market prices (e.g., Soltes (2009); Blankespoor et al. (2018)). The importance of the information production role of the financial press is more difficult to identify, also because it can take various forms and is consequently harder to isolate within financial press articles (Miller and Skinner (2015)). First, the press may assume an information aggregation function by linking information from various public sources, for example, by combining coverage of an earnings announcement with feedbacks from analysts during the conference call. Second, journalists may create original content using their individual analytical skills, filtering out critical items from a vast array of disclosed items, or by exploiting their access to private information from firm executives or analysts (e.g., via interviews). Moreover, journalists can create value by interpreting the disclosed information and by writing an article that is (more) understandable than the original disclosure.

In this paper, I examine the supply of accounting information. I use the release of the quarterly financial statements as the accounting event of interest and focus on the pertinent earnings announcements. The choice of a recurring event renders the results more generalizable to the principle work of business journalists, as the supply is independent of individual cases such as fraud. I choose three different news outlets: the Wall Street Journal (WSJ) as a financial newspaper and the New York Times (NYT) and USA Today (USAT) as general-interest newspapers. I analyze quarterly earnings announcements and corresponding articles of S&P 500 firms that are a permanent constituent of the index between Q1-2013 and Q2-2016. Overall, I analyze 323 firms. For each of these firms, I retrieve the quarterly earnings announcement and match the pertinent financial press articles.

I address three questions of interest. First, I am interested in the factors associated with the decision of news outlets to provide coverage of the earnings announcement of a firm. Particularly, I examine whether the coverage decision differs for outlets with varying reader sophistication. To answer this question, I consider the coverage pattern of the sample firms in the three different news outlets in a first step. I find that about 56% of the sample firms receive coverage in a news outlet over the sample period. This result is mainly driven by the WSJ, as the USAT and NYT only cover about 15% or 11%, respectively. The coverage decision of the different outlets appears to be sticky. Building on prior literature such as Thompson et al. (1987) and Fang and Peress (2009), I shed light on the very firm characteristics that shape demand for financial press intermediation using a determinants model in a second step. The analysis reveals that coverage by the financial press is positively associated with firm fundamentals like the size. Additionally, I find a positive association between the coverage and some of my proxies for the information environment, like analyst coverage.

On the other hand, I find no association with free float and Google searches, which are proxies of the general retail investor information demand. Also, my proxies for negative information contained in the earnings announcements yield mixed results. The coverage determinants do not vary substantially for the three different outlets, though it seems that general-interest newspapers favor more controversial firms. Taken together, the information environment of firms appears to be the primary driver of financial press coverage.

My second question pertains to the content of press articles, i.e., the nature and amount of information that they convey. Mostly, I am interested in the specific types and respective weight of information that the press decides to publish. To provide evidence on this matter, I examine the supply of information in financial press articles on the level of quantitative items and assign them to different categories. A systematic analysis of the content reveals that the outlets include an average of 17.7 items per article. When comparing the three different outlets, I find that specialized newspapers provide more quantitative items (20.5 vs. 16.0 and 16.7), which give readers more insights into the current and future development of a firm. In line with the prominent role of performance metrics, most accounting data that journalists decide to report are measures of profit and revenues. I include firm and analysts quotes into my analysis to assess how private communication may affect the information supply in financial press articles, as this supply is most likely based on information production. I find that journalists regularly include quotes in their articles and that about 75% of executive quotes in articles are based on information production. The NYT and USAT include more quotes from analysts and firm executives than the WSJ.

Finally, I am also interested in a potential production role of the press and therefore analyze to what extent press articles covering accounting events solely rely on firm items (i.e., disseminate information) or glean additional information from other sources (i.e., produce information), thereby providing a potential value-added to readers. To that end, I reconcile the quantitative items of the articles with quarterly firm disclosure. I find that journalists systematically include financial items and information gleaned from sources other than the earnings announcement, e.g., from the conference call or analysts into their articles. This result is consistent with press articles not only disseminating firm disclosures but also providing extra value by creating additional information. Mergers and acquisitions (89.5%) and expenses (37.4%) are among the categories for which news outlets supply most production items. Besides, I examine firm characteristics that are associated with additional information supply and therefore estimate a determinants model on the article level. Probit regressions reveal that the supply of this additional intermediation service increases with firm size and share turnover but is also markedly pronounced when retail investor information demand varies in the month leading up to the earnings announcement.

Overall, my analyses reveal that the press assumes its dissemination function by channeling information from earnings announcements, in particular on firm performance, to market participants. More importantly, I show that journalists also assume an information production role by combining information from various other sources and including expert quotes, providing potential transaction cost savings to their readers. Both the demand for dissemination and the supply of additional content increase with measures of the firm's information environment, such as size, which also reflects the notion of 'public interest'.

I contribute to the existing literature by providing an explorative supply-side perspective on the role of the financial press as an information intermediary. Generally, we know too little about whether and how the financial press assumes an information production function. Guest (2018) finds that articles add more value when they include more information that is original (i.e., uses a different wording than the earnings announcement), which suggests information production by the press. This finding contrasts Drake et al. (2014) who find that financial press coverage mitigates cash flow mispricing, but that this effect can be attributed only to news flashes disseminating news. My paper adds to this literature in at least two ways. First, to my knowledge, I am the first to provide detailed, content analysis-based evidence on the coverage decisions of the financial press, and on the choice of information items to be reported. Hence, I provide insights into the supply of intermediation activities by the press to the market. Second, my analysis of press article content enables me to discern the information dissemination role from the information production role of the financial press on a granular level. This way, my evidence helps us better understand how the press assumes its dissemination and production role and how journalists might create value through either role. My results complement Call et al. (2018), who offer insights into the information intermediary role of the press from the perspective of journalists.

The remainder of the paper is structured as follows: Section 2.2 elaborates on the theoretical role of the financial press as an information intermediary and prior literature.

Sections 2.3 to 2.6 present research design choices and findings from the analyses, which speak to my three research questions outlined above. Section 2.7 concludes.

2.2 Background and related research

Beaver (1998, p. 10) defines information intermediaries „an industry whose factors of production include financial information and other types of data and whose product is analysis and interpretation". The financial press is traditionally the information intermediary with the largest audience. Nevertheless, the capital market features various other information intermediaries such as financial analysts. Prior literature such as Drake et al. (2014) and Guest (2018)distinguishes two potential sources of value generation for the financial press, which I summarize in Figure 2.1: the dissemination of information (Information set I), and the production of information (Information set II and III).

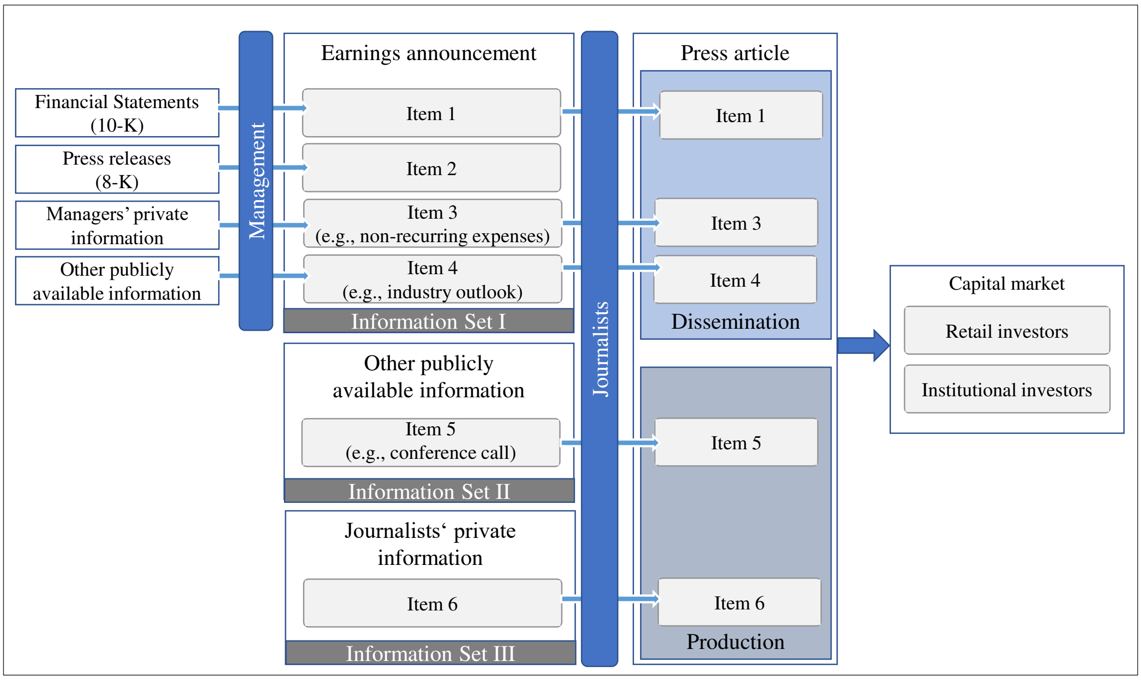


Figure 2.1: The financial press as an information intermediary
This figure shows the role of the financial press as an information intermediary through information dissemination (Information set I) and information production (Information set II and III). For both functions, the journalists collect, reduce, and weight the information available in each information set and base their articles on it. Subsequently, the articles become an information source for retail and institutional investors.

Generally, news outlets collect information from the three different information sets and make it available to their readers through their articles. Information set I features items from firm disclosure, for example, the earnings announcement press release (8-K) or financial statements. As these items are publicly available, Information set I relates to information dissemination. Information set II features other publicly available information, like information stated in conference calls, while Information set III includes journalists' private information, e.g., based on interviews with firm executives. Information from set II and III relate to information production because this kind of information expands the mere rebroadcasting of items. The journalists collect all items from the various information sets, decide which of them to include in the article, and how prominent they feature each item in the article. This collection and aggregation save information processing costs for investors.

There is ample evidence in the literature supporting the dissemination role of the financial press. For example, Peress (2014), Engelberg and Parsons (2011), Bushee et al. (2010), and Blankespoor et al. (2018) show a reduction in information asymmetry and a facilitated price discovery when the financial press covers a firm's earnings announcement. Supporting the information production role, Miller (2006) documents abnormal returns upon the publication of press articles discovering fraud cases. Miller (2006), alongside Core et al. (2008), further introduces a monitoring role of the financial press. The broad audience and the credibility of the financial press enhance the investor reaction to both disseminated and produced information (e.g., Kothari et al. (2009)). Based on these results, market reactions following press articles about earnings announcements could derive from either information production or dissemination.

In the context of earnings announcement related articles, Guest (2018) stresses the production role. He finds higher trading volume and more efficient price responses for articles with more production. His measure of information production focuses on the textual similarity between the earnings announcement press release and the press article. Based on this production measure, there are two ways in which journalists can create additional value to meet investor demand. The journalists can interpret the results and rephrase it in a way that is more understandable for (retail) investors. The more they rephrase the press release, the smaller the similarity between the two text documents. Alternatively, the journalists might include items from Information set II, or III in their articles, which results in a wording different to the earnings announcement. The proxy of Guest (2018) cannot discriminate between these two channels.

Contrasting the evidence outlined above, Drake et al. (2014) provide evidence that attributes the value of financial press articles to information dissemination. Their results show that press coverage of earnings announcements decreases cash flow mispricing. Yet these results only hold for newsflashes, which merely disseminate information. The authors fail to find a significant association for full articles, which might feature some degree of information production. One explanation for the contrasting findings might be that Guest (2018) focuses on the WSJ, a major financial newspaper, while Drake et al. (2014) examine a broader set of sources. Thus, the information production of different newspapers may vary with their degree of specialization. This difference might emanate from the demand of the readers the news outlets want to target. If news outlets were aware of investor demand in their target group, they might be more likely to write an article about a firm. Supporting this argument, Schütt (2019) finds evidence of market segmentation with differences in investor optimism and financial sophistication among various news outlets.

My paper contributes to the conflicting literature about the value of financial press articles. I conduct an exploratory analysis, which allows me to clearly distinguish between the information production and information dissemination within financial press articles. Thus, I show how the press assumes its two roles and how journalists might create value through either one of them. To account for diverging investor belief and sophistication, which leans them towards a particular newspaper, I examine the information supply of three different outlets. I believe that insights into coverage with respect to supply and content decisions of newspapers are important to better understand the factors that shape demand for press articles, and hence the channels of value creation. In line with the three questions of interest outlined in the introduction, I aim to (i) infer factors that shape the decision to supply intermediation and how they might differ between outlets with different target groups, to (ii) glean insights into the nature and relative proportion of information provided, and (iii) to understand whether and on what occasion journalists decide to not only disseminate information but also to assume an information production role.

2.3 Sample

I analyze S&P 500 firms between Q1-2013 and Q2-2016. I examine earnings announcements because they are scheduled in advance, can contain both positive and negative news, and investor information demand increases up to the event (Drake et al. (2012)). Additionally, they are independent of individual cases such as accounting fraud.

Following Hillert et al. (2014) and Fang and Peress (2009), I analyze articles covering quarterly earnings announcements published in the Wall Street Journal (WSJ), USA Today (USAT), and The New York Times (NYT). For a sample ending in 2002, Fang and

Peress (2009) find an overlap of about 75% between the different sources. Therefore, I can exploit the potential variation in coverage between the different outlets. While the WSJ is a specialized financial newspaper, the other two outlets are general-interest newspapers, thus I can analyze whether the information supply varies with outlet sophistication.

Table 2.1: Sample description

Panel A: Sample selection			
	Firms	Quarters	Firm-quarters
S&P 500 firms between 2013 and 2016	500	14	7,000
- Name conflicts Ticker Google	42	14	588
- Missing data	51	14	714
- balanced Panel	84	14	1,176
Balanced Sample coverage analysis	**323**	**14**	**4,522**

Panel B: Sample selection random sample		
	Firms	Firm-quarters
Balanced Sample coverage analysis	323	4,522
-Firms without press coverage in any quarter	49	686
-Firms without press coverage in some quarter	0	1,306
Potential Articles	274	2,530
Random article sample for content analysis	**134**	**420**

Panel C: Sample distribution per industry			
	Firms	Firm-quarters	Percent
(1) Consumer Non-Durables	24	336	7.43%
(2) Consumer Durables	7	98	2.17%
(3) Manufacturing	24	336	7.43%
(4) Oil, Gas, and Coal Extraction (Energy)	18	252	5.57%
(5) Chemicals and Allied Products	14	196	4.33%
(6) Business Equipment	49	686	15.17%
(7) Telephone and Television Transmission	10	140	3.10%
(8) Utilities	24	336	7.43%
(9) Wholesale, Retail, and Some Services	32	448	9.91%
(10) Healthcare, Medical Equipment	25	350	7.74%
(11) Finance	58	812	17.96%
(12) Other (e.g., Hotels, Entertainment)	38	532	11.76%
Total	*323*	*4,522*	*100.00%*

Notes: This table provides details on the sample selection process for the coverage (Panel A), the content analysis (Panel B), and sample distribution (Panel C).

Following Drake et al. (2014), I use RavenPack News Analytics to identify financial press articles. I only include full articles in my sample, which the news outlets publish in a seven-day window starting on the day before the earnings announcement day and ending five days afterward. To ensure that articles only deal with the sample firms, I exclude articles with a Ravenpack relevance score below 90[3]. If a firm is featured in an outlet multiple times during the seven-day window, I keep the article closest to the earnings announcement day for the analysis.

I use a balanced panel and therefore exclude all firms that are not a permanent constituent of the S&P 500 are have missing data in some quarters. Overall, I end up with 323 sample firms and 4,522 quarter-firm observations. Table 2.1, Panel A provides an overview of the sample selection process of the coverage sample, while Panel C provides information about the Fama and French 12-industry distribution of the sample.

2.4 Financial press coverage

My first research question relates to the factors associated with the decision of the financial press to provide coverage of a major accounting event (quarterly earnings announcement) for a firm. To provide evidence relating to those factors, I employ two different analyses. The first one provides descriptive evidence about the coverage distribution of the sample firms in the three outlets. For example, I am interested in how coverage varies between the different quarters and how sticky the coverage decision is on the news outlet level. This way, I can provide first evidence about potential differences between the various news outlets. The second analysis estimates a probit model to provide evidence which firm characteristics increase the likelihood of press coverage about the earnings announcement. In addition to the full model, I estimate the regressions for all news outlets separately, to provide further evidence about differences in the coverage decisions of the three outlets. Additionally, I investigate whether the supply is associated with retail investor demand. For these two analyses, I use the balanced sample outlined in Table 2.1 (Panel A).

2.4.1 Descriptive Analysis

In a first step, I am interested in the coverage distribution with regard to the weekday of the earnings announcement. In line with prior results (e.g., Tsileponis et al. (2020)), I find that coverage is most pronounced on the day of the earnings announcement (82.4%). Most

---

[3] For any news story that mentions an entity, RavenPack provides a relevance score between 0-100 that indicates how strongly related the entity is to the underlying news story, with higher values indicating greater relevance. A score of 0 means the entity was passively mentioned while a score of 100 means the entity was prominent in the news story. Usually, a relevance value of at least 90 indicates that the entity is referenced in the main title or headline of the news item, while lower values indicate references further in the story body.

of the sample firms report their earnings between Tuesday and Thursday. The percentage of firms covered does not differ substantially on these days (around 56% respectively). The likelihood of being covered is lowest on Monday and highest on Friday, though considerably fewer firms report their earnings on these days (Table 2.2, Panel A).

Table 2.2: Press Coverage of earnings announcements

Panel A: Press coverage of sample firms over earnings announcement weekdays

Weekday	Total	Without Coverage		With Coverage	
Monday	293	153	52.2%	140	47.8%
Tuesday	1,044	461	44.2%	583	55.8%
Wednesday	1,182	513	43.4%	669	56.6%
Thursday	1,576	700	44.4%	876	55.6%
Friday	422	164	38.9%	258	61.1%
Saturday	4	0	0.0%	4	100.0%
Sunday	1	1	100.0%	0	0.0%
Total	*4,522*	*1,992*	*44.1%*	*2,530*	*55.9%*

Panel B: Press coverage of sample firms over the sample period

Year/Quarter	All Papers	WSJ	USAT	NYT
2013q1	0.33	0.30	0.11	0.09
2013q2	0.33	0.30	0.12	0.09
2013q3	0.46	0.43	0.13	0.09
2013q4	0.59	0.59	0.13	0.08
2014q1	0.60	0.59	0.14	0.12
2014q2	0.64	0.63	0.14	0.12
2014q3	0.65	0.64	0.15	0.13
2014q4	0.64	0.63	0.19	0.14
2015q1	0.60	0.56	0.15	0.13
2015q2	0.60	0.60	0.19	0.12
2015q3	0.59	0.58	0.17	0.13
2015q4	0.64	0.63	0.17	0.14
2016q1	0.59	0.57	0.16	0.13
2016q2	0.57	0.56	0.12	0.11
Total	*0.56*	*0.54*	*0.15*	*0.11*

Notes: This table provides information on the press coverage of earnings announcements. Panel A shows the coverage regarding the weekday of the earnings announcement. Panel B shows the median coverage of all 323 sample firms for all sample quarters in total and by news outlet.

In a second step, I examine the average coverage of the 323 sample firms in the three outlets to provide preliminary evidence about potential differences between them. Table 2.2, Panel B shows the coverage for each news outlet over the sample period. I find that the news outlets feature about 56% of the sample S&P500 firms over time. The coverage varies substantially between the outlets. While the WSJ covers 54% of the firms on average, the USAT and NYT only pick up 15% and 11% of earnings announcements, respectively. For the USAT and WSJ, these numbers are comparable to Fang and Peress (2009), whereas the NYT coverage is considerably lower in my sample[4]. In line with Guest (2018), I find an increase in coverage during 2013, when the WSJ merged its news desk with the Dow Jones Newswire[5].

Out of the 323 sample firms, 49 are not featured in any of the three outlets during the sample period, whereas 48 of them receive coverage in every quarter of the sample in at least one outlet. Interestingly, 27 firms receive constant coverage in one specific news outlet after the initial coverage during the sample period. This effect might be comparable to analyst coverage (Clement (1999)): Once an analyst or a journalist decides to follow a firm, she keeps publishing reports about that firm and uses firm-specific knowledge/experience to provide more accurate information. However, the WSJ mainly drives the coverage distribution. While the WSJ keeps publishing articles about a firm once the coverage started, the it is less constant for specific firms in the other two outlets. Figure 2.2 provides an overview of the firm coverage distribution of the sample based on 14 quarters.

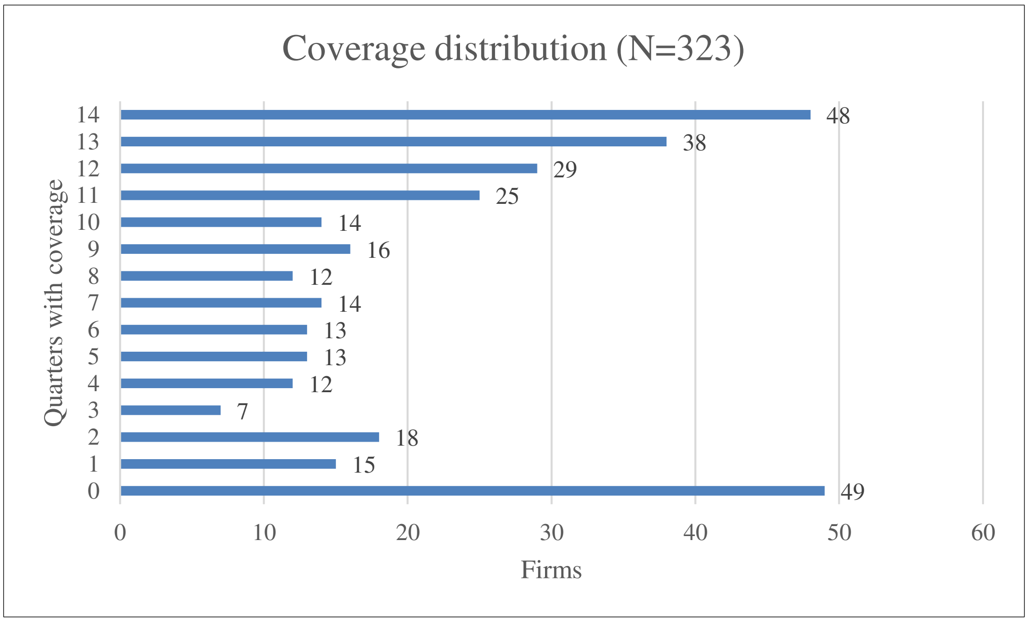


Figure 2.2: Coverage distribution of sample firms
This figure shows the number of quarters in which one of the three new outlets (WSJ, NYT, USAT) features the firms (N Quarters=14, N Firms =323). 48 firms are covered in all quarters, while 49 firms are never included in a press publication.

[4] The sample does not include articles below a relevance score of 90. This means that a firm could be mentioned in either NYT or USAT in an article covering various firms, in which it would be of subordinate importance (e.g., in articles that summarize industry earnings announcements).
[5] https://www.politico.com/blogs/media/2013/06/wsj-dow-jones-to-restructure-167015

2.4.2 Determinant Analysis

2.4.2.1 Design

To provide insights into the coverage decision of the outlets, I examine which firm characteristics increase the likelihood of coverage and especially whether these characteristics differ for the news outlets. I estimate a probit determinants model using the coverage (indicator variable set to one if a firm is covered in the current quarter and zero otherwise) as the dependent variable. My determinants model aims to capture firm characteristics that potentially shape demand for financial intermediation via newspaper articles. To that end, I distinguish three sets of test variables.

My first vector of variables relates to the general information environment of the firm. While institutional investors have access to costly information intermediaries, e.g., financial analysts, retail investors rely on less costly information intermediaries like newspapers. I include the free float of total shares outstanding (*Freefloat*) (Drake et al. (2014)) as a proxy for the ownership structure. My descriptive evidence suggests that journalists, similar to analysts, continuously cover a firm once they initiated coverage. Therefore, I include the press coverage in the previous quarter (*Cov Q-1*) as an additional variable (Drake et al. (2014)). Journalists and news outlets have limited resources. Consequently, they might be less able to devote resources to an article when multiple firms are announcing their earnings on the same day. Therefore, I use a median split to create the variable *BusyDay*, which captures days on which more firms than usual announce their earnings. According to Call et al. (2018), the main attribute that influences the coverage decision of a journalist is the question of whether a firm is controversial. Prior research about the coverage decisions of news outlets has so far neglected this argument in the analyses. To capture the controversy of a firm leading up to the earnings announcement, I use the Ravenpack aggregated event sentiment score (*AES*), which captures the ratio of positive to negative news stories within a 91-day window. The score does not consider neutral articles. A score of 0 indicates that all articles within the window are negative, while a score of 1 indicates all articles are positive.

Financial analysts are among the essential information intermediaries. For some investors, analysts represent an important alternative information source to the press. So far, the literature has not clearly established whether the two information intermediaries are complementary or substitutes (Miller and Skinner (2015)). Hillert et al. (2014) and Fang and Peress (2009) provide results indicating that the two intermediaries are substitutes. On the other hand, Call et al. (2018) provide evidence that journalists include analyst information in their articles, suggesting that they consider both intermediaries to be complementary. In line with this argument, the results of Guest and Jaewoo (2019) indicate that financial press coverage decreases when fewer analysts are available as a potential source to journalists. Besides, Rees et al. (2015) argue that financial press coverage is favorable for the career prospects of analysts. To provide evidence on this ongoing debate, I include the analyst following into the determinant model. I measure *AF* as the natural logarithm of one plus the number of EPS estimates.

As I am interested in the supply-demand relation of news outlets and their target groups, I include a proxy for investor information demand in my model. If news outlets were aware of retail investor demand, they might be more likely to publish an article about the earnings announcement. Da et al. (2011) use the Google search volume index (*SVI*) based on firm tickers as a proxy for (retail) investor demand for information about a firm. Using the SVI data, Drake et al. (2012) find that Google searches are positively associated with media coverage.[6] Following these papers, I use the Google *SVI* as a proxy for the retail investor information demand leading up to the earnings announcement. In the various analyses, I use two different versions of this variable. The first one is the natural logarithm of the mean SVI during the 28 days leading up to the earnings announcement (*SVI_mean*). The second one uses the standard deviation of SVI during the same time (*SVI_sd*). The variables based on the Google data are of special interest to me because they capture whether demand is associated with supply.

My second vector of variables captures the information contained in the earnings announcement. Broadly, I conjecture that demand for newspaper coverage increases with the novelty of the information, e.g., the earnings surprise (Engelberg and Parsons (2011)). I use the analyst forecast error (*AFE*) as a measure for earnings surprise. I calculate *AFE* as the difference between the actual earnings and the last analyst forecast before the earnings announcement and scale the difference by the stock price at the quarter-end (Bushee et al. (2010)). Additionally, prior literature shows that journalists are more likely to cover negative firm events, as they anticipate increased investor demand. I use the indicator

[6] The authors examine S&P 500 firms between 2005 and 2008 and find an increase in Google searches in the two weeks leading up to the earnings announcement with a spike on the earnings announcement day. They use the media data of Soltes (2009), which includes the WSJ, New York Times, The Washington Post and USA Today.

variable *Loss_firm* to proxy for the disclosure of bad news (i.e., reported loss) in a specific quarter. As an alternative proxy for negative firm events, I include the indicator variable *AF_miss,* which captures if firms meet the analyst forecast. Call et al. (2018) argue that journalists compare the reported earnings to the analyst forecast. Therefore, they might be more likely to cover an earnings announcement when they can write about a missed forecast.

My third vector includes variables capturing the fundamentals of the sample firms. In general, larger and potentially growing firms are more interesting for the market participants. Additionally, less sophisticated news outlets might mainly feature large firms, as their readers might be less familiar with smaller firms. I measure *Size* as the natural logarithm of market capitalization at the fiscal quarter-end. I include the book-to-market ratio (*BTM*) to capture the growth potential of the firm (Bushee et al. (2010)). I use the return volatility during the previous quarter beginning on the day of the last earnings announcement and ending one day before the current earnings announcement (V*olatility)* to proxy for the risk of a firm. I capture the liquidity of the firm's stock using the mean share turnover (*Turnover*) over the previous quarter (Fang and Peress (2009)). For the same period, I calculate buy and hold returns (*BHR*) to include the stock performance of a firm into the model. Call et al. (2018) show that journalists take the stock performance (weak or strong) of a firm into consideration when deciding whether to cover a firm or not.

I obtain weekly Search Volume Index (SVI) data from Google Trends. Following prior literature (e.g., Drake et al. (2012); Da et al. (2011); Madsen and Niessner (2019)), I use company tickers to collect the search data. I exclude 42 firms with tickers with potential alternative meanings (e.g., CAT) and one firm because it does not have enough search requests for Google to provide SVI data[7]. I use firm fundamentals data from Compustat and market data from CSRP. To include analysts as a different outlet of information intermediaries, I use I/B/E/S data. I obtain ownership data from Datastream.

I use the Fama and French 12 industry classification to control for industries that generally interest the public, e.g., banks. To deal with outliers, I winsorized *Size* and *AFE* at the 1% level. Additionally, I include quarter-year fixed effects and cluster standard errors on the firm level. I provide variable definitions in Table 2.A1 (Appendix) and the summary

[7] Google SVI data ranges from 0 to 100. If a term does not have enough searches, Google does not display results for the term.

statistics in Table 2.3. Table 2.A2 in the appendix shows the correlation between the various variables.

Table 2.3: Summary statistics (Balanced coverage analysis)

Variable	mean	sd	median	Min	max	Count
Coverage All Papers	0.56	0.50	1.00	0.00	1.00	4,522
WSJ Coverage	0.54	0.50	1.00	0.00	1.00	4,522
USAT Coverage	0.15	0.35	0.00	0.00	1.00	4,522
NYT Coverage	0.11	0.32	0.00	0.00	1.00	4,522
Cov Q-1	0.55	0.50	1.00	0.00	1.00	4,522
NYT Cov Q-1	0.12	0.32	0.00	0.00	1.00	4,522
USAT Cov Q-1	0.14	0.35	0.00	0.00	1.00	4,522
WSJ Cov Q-1	0.53	0.50	1.00	0.00	1.00	4,522
AES	0.65	0.15	0.65	0.15	1.00	4,522
BusyDay	0.48	0.50	0.00	0.00	1.00	4,522
Size	10.03	0.99	9.88	8.05	12.68	4,522
Freefloat	0.91	0.10	0.93	0.48	1.00	4,522
BTM	0.40	0.30	0.32	-0.92	2.18	4,522
Loss firm	0.03	0.16	0.00	0.00	1.00	4,522
SVI_mean	3.73	0.62	3.88	0.00	4.61	4,522
SVI_sd	6.05	5.26	4.51	0.00	44.40	4,522
Turnover	8.35	5.51	7.01	0.01	83.58	4,522
Volatility	0.01	0.01	0.01	0.01	0.10	4,522
AF	2.97	0.45	3.04	0.69	4.04	4,522
AFE	-0.04	0.02	-0.04	-0.14	0.02	4,522
AF_miss	0.02	0.13	0.00	0.00	1.00	4,522
BHR	0.04	0.12	0.04	-0.71	1.70	4,521

Notes: This table provides summary statistics for the variables of the coverage analysis. Sample selection is defined in Table 2.1. All variables are defined in Table 2.A1 (Appendix). Variables comprise Coverage (outlets combined and separately), prior coverage (outlets combined and separately), AES (press sentiment score), BusyDay dummy, Size, Freefloat, BTM (Book-to-market), loss firm dummy, SVI_mean (Google search mean), SVI_sd (Google search standard deviation), Turnover, volatility, AF (number of analyst forecasts), AFE (analyst forecast error), AF_miss (dummy for missed forecast), and BHR (buy-and-hold return).

2.4.2.2 Results

Table 2.4 reports findings for estimating the determinants model based on the variables outlined above. Column (1) shows the results for all newspapers.[8] Columns (5) and (6) show various model specifications using the same dependent variable but different test variables to assess the robustness of the results. The coefficient estimates for *Size* are

[8] The results remain mainly unchanged when estimating the determinants model without clustering the standard error or including fixed effects.

highly significant (1% level) in all model specifications. The positive coefficients suggest that a firm is more likely featured in the financial press when it has a high market capitalization. This result is in line with readers' information demand, as retail investors, who particularly rely on information provided in the financial press (Kothari et al. (2009)), are more likely to invest in well-known firms and thus want more information about them (Barber and Odean (2008)). Share turnover during the quarter is also positive and significant (1% level), indicating that news outlets provide more information about highly traded stock and therefore cater to the demand of their readers.

SVI_sd and *SVI_mean*, which capture the investor information demand leading up to the earnings announcement, yield no significant results in any model specification. This result suggests that news outlets do not adapt their supply to a short-term increase in demand. Coefficient estimates for the number of analyst EPS forecasts (*AF*) are significant (1% level) and obtain positive signs in all models. This indicates that firms, which attract attention by analysts, are also more likely to receive public attention in a newspaper. To that end, there seems to be a certain overlap between the target groups of analysts and the news outlets, which suggests that the press and analysts complement rather than substitute each other. Prior press coverage (*Cov Q-1*) is positive and highly significant (1% level) in all models, confirming the descriptive results that journalists keep writing about a firm once they started covering. The *BusyDay* dummy is significantly negative, which suggests that press coverage is less likely on days with many earnings announcements, which is in line with time and newspaper space constraints. The variable for the ownership structure (*Freefloat*) remains insignificant in all model specifications. The results for *AES*, which is a proxy for the degree of controversy surrounding the firm, are positive and significant at a 1% level in all model specifications. This suggests that coverage is more likely when prior articles are more favorable.

The coefficient of the analyst forecast error (*AFE*) is negative and significant on a 1% level in all model specifications. Additionally, the loss firm indicator variable has a positive association with press coverage on the 5% level. These results are consistent with journalists supplying more information when the earnings announcement conveys bad news. The positive association between volatility and press coverage (column (5) further suggests that high-risk firms are more likely featured in a newspaper.

Table 2.4: Coverage Analysis

	(1) Coverage all papers	(2) Coverage NYT	(3) Coverage USAT	(4) Coverage WSJ	(5) Coverage all papers	(6) Coverage all papers
Size	0.448***	0.810***	0.681***	0.398***	0.405***	0.357***
	(8.06)	(11.25)	(10.66)	(7.49)	(8.06)	(7.28)
Freefloat	-0.088	-0.838	-1.166**	-0.064	-0.001	-0.163
	(-0.24)	(-1.44)	(-2.23)	(-0.18)	(-0.00)	(-0.45)
BTM	0.065	0.437*	0.195	0.031	0.112	0.142
	(0.49)	(1.68)	(0.92)	(0.24)	(0.82)	(1.07)
AF	0.285***	0.163	0.214*	0.303***	0.332***	0.353***
	(3.93)	(1.07)	(1.68)	(4.19)	(4.67)	(4.86)
AFE	-4.332***	-9.069***	-5.976**	-3.573**	-3.548**	-3.907***
	(-2.91)	(-2.91)	(-2.25)	(-2.51)	(-2.05)	(-2.61)
SVI_sd	0.002	0.003	0.003	-0.002		
	(0.46)	(0.32)	(0.40)	(-0.37)		
Turnover	0.031***	0.036***	0.035***	0.029***		
	(3.82)	(3.22)	(3.68)	(3.80)		
Cov Q-1	1.337***	0.855***	0.874***	1.311***	1.336***	1.374***
	(15.71)	(5.64)	(7.07)	(15.84)	(16.03)	(16.45)
AES	2.240***	-0.327	-0.549**	2.339***	2.197***	2.162***
	(10.65)	(-1.24)	(-2.16)	(11.43)	(10.62)	(10.26)
Loss firm	0.384**	0.146	-0.126	0.428***		0.536***
	(2.34)	(0.54)	(-0.51)	(2.64)		(3.35)
BusyDay	-0.106*	-0.223**	-0.080	-0.112**	-0.113**	-0.114**
	(-1.91)	(-2.50)	(-0.89)	(-2.02)	(-2.02)	(-2.06)
SVI_mean					0.079	0.061
					(1.37)	(1.10)
Volatility					26.519***	
					(5.01)	
AF_miss					0.061	
					(0.31)	
BHR						-0.264
						(-1.29)
Pseudo R^2	0.424	0.376	0.342	0.415	0.423	0.419
N	4,522	4,522	4,522	4,522	4,522	4,521
Fixed Effect	Quarter, Industry	Quarter, Industry	Quarter, Industry	Quarter, Industry	Quarter, Industry	Quarter, Industry

Notes: This table shows the results of a probit regression with the coverage as the dependent variable. In columns (1), (5), and (6) the dependent variable refers to all news outlets, while columns (2) to (4) estimate the model for the three different outlets outlined in the column's headline. All variables are defined in Table 2.A1 (Appendix). Variables comprise Coverage (outlets combined and separately), prior coverage (outlets combined and separately), AES (press sentiment score), BusyDay dummy, Size, Freefloat, BTM (Book-to-market), loss firm dummy, SVI_mean (Google search mean, SVI_sd (Google search standard deviation), Turnover, volatility, AF (number of analyst forecasts), AFE (analyst forecast error), AF_miss (dummy for missed forecast), and BHR (buy-and-hold return). ***, **, * indicate statistical significance at the 1%, 5%, and 10% level, respectively, using two-tailed tests and standard errors clustered at the firm level. (N firms=323). Maximal VIF in all models is 3.38.

Table 2.4, columns (2) to (4) show the results for estimating the model for the three outlets separately. The results for *Size*, *Turnover*, *Cov Q-1* (in the respective newspaper), and *BusyDay* remain mostly unchanged. Interestingly, the coefficient estimate for the number of analyst forecasts is only highly significant for the WSJ. This suggests that analyst coverage is increasing the coverage of a specialized newspaper, while it is less likely for newspapers of general-interest. The coefficient estimate for *AES* is positive and slightly significant in the full models. This result holds for the WSJ, but the sign switches for the other two outlets. The negative coefficient estimate is significant at the 5% level for the USAT. Therefore, the results suggest that the two general-interest newspapers of my sample are more likely to cover an earnings announcement when prior press coverage is more negative. This is in line with the results of Call et al. (2018) that journalists feature more controversial firms, which speaks more to the entertainment role of the press rather than disseminating or producing information (Jensen (1979)). While the results for *AFE* remain mainly unchanged, the loss firm indicator variable only has a positive association with the WSJ coverage. This hints that the WSJ is more likely to feature information about constrained firms that do not meet the market expectations. In contrast to the other outlets, USAT coverage is significantly negative associated with the number of free float shares.

Overall, my results suggest that the general firm fundamentals and the information environment of the firm are a vital factor for the decision of a news outlet to cover the earnings announcement. Additionally, there seems to be more coverage when the information conveyed in the earnings announcement is negative. While the results suggest that most firm characteristics drive coverage in all three outlets, the WSJ journalists react stronger to professional information sources, for example, analysts, whereas USAT and NYT journalists supply articles for more controversial firms, which might be more interesting to their targeted readers. These results are in line with the level of specification of the news outlets and complement those of Schütt (2019) regarding news market segmentation from a supply perspective.

2.5 Information supplied in financial press articles

2.5.1 Design

My second research question relates to the content of press articles, i.e., the information that journalists supply in their articles. I expect insights into the choice and weighting of content in press articles to be reflective of readers' demand for information and hence for intermediation services provided by the press. To provide evidence on the information supplied in the articles, I perform a content analysis aimed at collecting all quantitative items and all quotes. While the quantitative items represent hard information, which is value relevant for investors, the quotes are indicative of additional content and reflect possible private communication (i.e., information production; Call et al. (2018)). By performing a content analysis, I can provide detailed evidence about the information supply in articles, which is more granular compared to the textual similarity measure of Guest (2018) and complements the arguments Call et al. (2018) make in their survey paper about the information supply from the perspective of journalists. I identify 2,530 articles about the sample firms during the sample period. I use a random sample for each news outlet in each quarter containing 10 articles for the manual collection. In sum, I analyze 420 articles. The analyses are based on the balanced sample for the coverage analysis outlined in Table 2.1, Panel A. For the content analysis, I obtain the text of the articles from ProQuest and Nexis.

The content analysis focuses on three main areas of interest. First, I am interested in the general characteristics of the financial press articles and how they differ between the three outlets. To that end, I examine the length, the overall supply of quantitative items, and the information density in the different articles. As readers of a specialized newspaper might be more sophisticated, WSJ articles present the information differently than the more general outlets.

Second, I examine the time horizon of the quantitative items (current, forecast, prior). Most items likely refer to the quarter of the current earnings release. Nevertheless, it might be beneficial for readers if the journalists include items from the previous quarter or year to facilitate a comparison and, thus, the evaluation of the development of the firm. Additionally, it might be in line with the readers' demand if the articles feature management and/or analyst forecasts to compare the forecasted firm development with their own expectations.

Third, I categorize the quantitative items to examine the depth of the articles and to assess the level of supply of each news outlet. This way, I can provide evidence on how diversified journalists portray the firms, e.g., by reporting segment performance or disaggregated earnings components. While disaggregated earnings might be helpful to mitigate investor fixation on earning (Elliott et al. (2011)), a broader choice of firm items such as segment performance indicators might facilitate the investment decision for investors. I assign each quantitative item to a group: accounting data, non-accounting data, and external data. I then further divide the groups into different categories.

Table 2.5: Category examples

Panel A: Internal Accounting data	
Category	Examples
Balance Sheet	assets, provision, debt
Cash Flow	FCF, cash flow, operating cash flow
EB	operating earnings, core earnings, operating income
Expenses	impairment, expense, restructuring costs
GAAP	net income, earnings, EPS
Industry-specific	tenant retention, mortgage applications, display ads, corn volume
M&A	acquisition, takeover bid
Non-GAAP	adjusted earnings, adjusted profit, adjusted eps
Revenue	revenue, sales, gross margin
Segment	segment earnings, segment operating income, segment revenue
Share related	dividend, share buyback
Panel B: Internal Non-Accounting data	
Category	Examples
Employees	workforce, wages, bonus
Exchange rate	currency effect, currency hedge
Legal	settlement, litigation, lawsuit
Other	customers, stores, taxes, Brexit
Product related	product price, production, backlog
Panel C: External data	
Category	Examples
Analyst	revenue forecast, EPS forecast
Capital Market	share price, S&P500, FTSE
Competitor	sales, earnings, EPS
Industry	market share, federal funding, coal shipments
Macro Data	GDP, oil price
Panel D: Quotes	
Category	Examples
Internal Quote	CFO quote, CEO quote, Conference Call quote
Analyst Quote	Analyst quote
External Quote	Former employees, research

Notes: This table shows examples for the items in each of the categories accounting data (Panel A), non-accounting data (Panel B), external data (Panel C), and quotes (Panel D).

The accounting data group comprises categories like revenue, performance, and segment measures. There is an ongoing debate about the motive of managers to disclose non-GAAP figures (e.g., Leung and Veenman (2018)). Consequently, it is of interest which performance metrics journalists choose to include in their articles. I follow Hitz (2010) and divide the performance category into three subcategories: GAAP, EB, and non-GAAP. I define GAAP items as a bottom-line measure from the income statement and EB items as "earnings before" metrics, for example, EBIT or EBITDA. Non-GAAP metrics are adjusted performance metrics such as a recurring EBITDA. For these metrics, the management decides to eliminate expenses like restructuring costs. Therefore, additional information is needed for these metrics to be reconciled with the income statement.

The non-accounting data group contains information about employees or product-related items, while the external data group contains analyst forecasts, market prices, and macroeconomic figures such as GDP growth. I assign all quotes to my last group. I split this group into analyst, firm, and external quotes. For the firm and analyst quotes, I additionally analyze whether they emanate from public sources like a conference call (Information set II) or private communication like interviews (Information set III). While firm executives have incentives to talk to journalists to impact the reporting about their firms, analysts might draw career benefits from their public coverage (Rees et al. (2015)). Featuring quotes allows news outlets to include the perspective of users and preparers into the articles and signals access to background information to the readers. The inclusion of this kind of information is likely to produce valuable information for the readers. Table 2.5 gives an overview of the different categories in the groups and provides examples for each category.

2.5.2 Quantitative items: general article information

Overall, I collect 7,447 quantitative items and 546 quotes from the articles. Table 2.6 (Panel A) reports the general characteristics (mean lengths, items, and density) of all articles in total and for the different news outlets, while Figure 2.A3 (Appendix) shows the mean items per outlet for the different quarters to evaluate a time trend. Table 2.6 (Panel B) provides the results of t-tests comparing the mean items and mean length among the different outlets. The articles have a mean word count of 484.4. On average, articles of the NYT have more words than the ones of the other two outlets, while WSJ articles have a larger minimum of words. The difference between the NYT and the other two outlets is statically significant at a 1% level, whereas WSJ articles are also significantly longer than USAT articles (1% level) (Panel B). News outlets include a mean of 17.7 items in each article, which results in a mean of 0.04 items per word in each article. WSJ articles contain an average of 20.5 key items in each article, while the NYT (16.0) and the USAT (16.7) include significantly fewer items (significant on a 1% level). The difference between the two general-interest newspapers is not statistically significant (Table 2.6, Panel B). Overall, it seems like the specialized WSJ has a higher information density. As its target group is likely made of more sophisticated readers/investors, the WSJ might choose to supply more detailed information. This is in line with Schütt (2019), who suggests that news outlets cater to their specific target group.

Table 2.6: Quantitative items in financial press articles

Panel A: Quantitative items in financial press articles

		Total	NYT	USAT	WSJ
Word Count	Mean	484.4	612.9	374.4	466.0
	Min	35	87	35	172
	Max	1,285	1,285	773	1,029
Items	N	7,447	2,234	2,337	2,876
Items per article	Mean	17.7	16.0	16.7	20.5
	Min	1	1	3	3
	Max	53	41	42	53

Panel B: Outlet comparison

	N	Mean Items per article	Mean word count
NYT	140	16.0	612.9
WSJ	140	20.5	466.0
Diff		-4.59***	146.9***
USAT	140	16.7	374.4
WSJ	140	20.5	466.0
Diff		-3.8***	-91.6***
NYT	140	16.0	612.9
USAT	140	16.7	374.4
Diff		0.7	238.5***

Notes: This table shows quantitative items in financial press articles in total and by news outlet. Panel A reports the word count, items, and items per article in the sample news outlets. Panel B shows the results of t-tests comparing mean items and word counts of articles published in the different outlets. ***, **, * indicate statistical significance at the 1%, 5%, and 10% level, respectively. The analysis is based on 420 articles from three news outlets (140 each).

2.5.3 Quantitative items: time horizon

Table 2.7, Panel A shows the time horizon of the quantitative items. As expected, most items are based on the current quarter. In line with the results of Call et al. (2018), nearly all articles (N=410 out of 420) feature some information about a previous quarter, as journalists often use them as a comparison for the current results. Table 2.7, Panel B splits the items into different groups. Newspapers compare accounting items to the previous quarter more often, while they chose not to compare the non-accounting items. Table 2.7, Panel C shows the time horizon distribution by news outlet. The WSJ supplies relatively more forecasts than the other two outlets. For the accounting items, this larger number is mainly due to earnings and revenue forecasts. Additionally, the WSJ features much more analyst forecasts than the other two outlets (N=62 vs. N=16 and N=17). Therefore, readers can evaluate the firm performance based on expert input, which might be more in line with the information demand of a WSJ reader. Overall, the journalists focus on the new information presented in the earnings announcement, with the WSJ giving a more detailed outlook for future development.

Table 2.7: Time horizon of quantitative items

Panel A: Time horizon						
	N (Items)	%	N (Articles)	%		
Current FY	4,524	60.7%	420	100.0%		
Forecast	585	7.9%	224	53.3%		
Prior FY	2,338	31.4%	410	97.6%		
Total	*7,447*	*100.0%*	*420*	*100.0%*		
Panel B: Time horizon by group						
	External data		Accounting data		Non-Accounting data	
	N (Items)	%	N (Items)	%	N (Items)	%
Current FY	1,082	69.3%	3,117	57.7%	325	67.0%
Forecast	116	7.4%	384	7.1%	85	17.5%
Prior FY	364	23.3%	1,899	35.2%	75	15.5%
Total	*1,562*	*100.0%*	*5,400*	*100.0%*	*485*	*100.0%*
Panel C: Time horizon by news outlet						
	NYT		USAT		WSJ	
	N (Items)	%	N (Items)	%	N (Items)	%
Current FY	1,458	65.3%	1,443	61.7%	1,623	56.4%
Forecast	118	5.3%	146	6.2%	321	11.2%
Prior FY	658	29.5%	748	32.0%	932	32.4%
Total	*2,234*	*100.0%*	*2,337*	*100.0%*	*2,876*	*100.0%*

Notes: This table shows the time horizon distribution of quantitative items for all items (Panel A), by group (Panel B), and by news outlet (Panel C). The analysis is based on 420 articles from three news outlets (140 each). Percentage number refer to all items of the *total* number of items.

2.5.4 Quantitative items: categories

Table 2.8 (Panel A) summarizes the results of the content analysis for the different groups. Out of all items, 79% are based on internal (accounting and non-accounting) data. My results show that most items (5,400/72.5%) belong to the accounting data group. Nearly every article contains information about at least one item of this group. As performance and revenue are two important key figures for investors, it is not surprising that journalists often report items of these categories (Panel B). Together, the two categories comprise more than 50% of the accounting group (2,033 items/37.7% performance; 1,208 items/22.4% revenue). More than 90% of the articles (91.9% performance and 97.1% revenue) contain at least one item out of these two categories.

Another important category in the accounting data group are the segment items (1,464 items/ 27.1%). As the performance of segments can differ from the performance of a full firm, this information might be valuable to readers to assess the current and future performance of a firm in a more diversified way. The use of segment information varies between the outlets. While nearly 80% of WSJ articles provide this more detailed insight, the percentage of articles containing segment information decreases for the general-interest newspapers (NYT: 70.7%; USAT: 63.6%). For segment items, the difference between the WSJ and the general-interest outlets is significant at a 5% and 10% level, respectively.

While the two general outlets provide a comparable amount of performance items (603 and 571, respectively), the WSJ includes most of them (859) (Panel C). Of the three performance categories, journalists include the GAAP metrics most often in their articles. The frequency with which the general-interest newspapers report the three different performance metrics are comparable (GAAP: 92.9% vs. 93.6%; EB: 11.4% vs. 11.4%; non-GAAP: 27.9% vs. 36.4%), whereas the WSJ features EB performance measures in 20% and non-GAAP metrics items in more than half of their articles (52.1%). As Regulation G requires firms to present a comparable GAAP measure and an at least equal presentation of both metrics when disclosing a non-GAAP measure, performance items clearly show a supply choice by the news outlets. In general, the WSJ supplies more quantitative items that allow investors to differentiate transitory and persistent performance components, as it includes prior mergers and its effects as well as detailed information about expenses in the articles. This disaggregation of earnings might mitigate earnings fixation on behalf of the investors. However, the disclosure of non-GAAP measures by the WSJ might also undermine the requirement of equal presentation of Regulation G if the news outlet fails to report the corresponding GAAP metrics. 25% of journalists in the survey paper of Call et al. (2018)state that they are not likely to evaluate the difference between GAAP and non-GAAP metrics in their articles.

Table 2.8: Items in financial press articles

Panel A: All items

	Total				NYT				USAT				WSJ			
Group	Items	%	Art.	%	Items	%	Art.	%	Items	%	Art.	%	Items	%	Art.	%
Accounting	5,400	72.5%	417	99.3%	1,656	74.1%	138	98.6%	1,567	67.1%	139	99.3%	2,177	75.7%	140	100.0%
Non-Accounting	485	6.5%	316	75.2%	156	7.0%	106	75.7%	154	6.6%	114	81.4%	175	6.1%	96	68.6%
External	1,562	21.0%	400	95.2%	422	18.9%	129	92.1%	616	26.4%	137	97.9%	524	18.2%	134	95.7%
All items/Articles	*7,447*	*100.0%*	*420*	*100.0%*	*2,234*	*100.0%*	*140*	*100.0%*	*2,337*	*100.0%*	*140*	*100.0%*	*2,876*	*100.0%*	*140*	*100.0%*

Panel B: Accounting data

	Total				NYT				USAT				WSJ			
Category	Items	%	Art.	%	Items	%	Art.	%	Items	%	Art.	%	Items	%	Art.	%
Balance Sheet	33	0.6%	21	5.0%	18	1.1%	12	8.6%	9	0.6%	5	3.6%	6	0.3%	4	2.9%
Cash-Flow	35	0.6%	24	5.7%	16	1.0%	11	7.9%	2	0.1%	2	1.4%	17	0.8%	11	7.9%
Expenses	206	3.8%	109	26.0%	51	3.1%	28	20.0%	47	3.0%	30	21.4%	108	5.0%	51	36.4%
Industry-specific	253	4.7%	84	20.0%	71	4.3%	25	17.9%	43	2.7%	18	12.9%	139	6.4%	41	29.3%
M&A	65	1.2%	53	12.6%	13	0.8%	11	7.9%	13	0.8%	12	8.6%	39	1.8%	30	21.4%
Performance	2,033	37.6%	408	97.1%	603	36.4%	135	96.4%	571	36.4%	134	95.7%	859	39.5%	139	99.3%
Revenue	1,208	22.4%	386	91.9%	378	22.8%	126	90.0%	360	23.0%	127	90.7%	470	21.6%	133	95.0%
Segment	1,464	27.1%	296	70.5%	474	28.6%	99	70.7%	488	31.1%	89	63.6%	502	23.1%	108	77.1%
Share related	103	1.9%	56	13.3%	32	1.9%	20	14.3%	34	2.2%	18	12.9%	37	1.7%	18	12.9%
All items/Artciles	*5,400*	*100.0%*	*420*	*100.0%*	*1,656*	*100.0%*	*140*	*100.0%*	*1,567*	*100.0%*	*140*	*100.0%*	*2,177*	*100.0%*	*140*	*100.0%*

Panel C: Performance Items

	Total				NYT				USAT				WSJ			
Category	Items	%	Art.	%	Items	%	Art.	%	Items	%	Art.	%	Items	%	Art.	%
GAAP	1,617	79.5%	397	94.5%	509	84.4%	130	92.9%	460	80.6%	131	93.6%	648	75.4%	136	97.1%
EB	132	6.5%	60	14.3%	37	6.1%	16	11.4%	28	4.9%	16	11.4%	67	7.8%	28	20.0%
Non-GAAP	284	14.0%	163	38.8%	57	9.5%	39	27.9%	83	14.5%	51	36.4%	144	16.8%	73	52.1%
All items/Articles	*2,033*	*100.0%*	*420*	*100.0%*	*603*	*100.0%*	*140*	*100.0%*	*571*	*100.0%*	*140*	*100.0%*	*859*	*100.0%*	*140*	*100.0%*

Table 2.8: Items in financial press articles (continued)

Panel D: Non-Accounting data

	Total				NYT				USAT				WSJ			
Category	Items	%	Art.	%	Items	%	Art.	%	Items	%	Art.	%	Items	%	Art.	%
Employees	102	21.0%	58	13.8%	35	22.4%	21	15.0%	35	22.7%	20	14.3%	32	18.3%	17	12.1%
Exchange rate	37	7.6%	26	6.2%	5	3.2%	4	2.9%	7	4.5%	5	3.6%	25	14.3%	17	12.1%
Legal	11	2.3%	7	1.7%	10	6.4%	6	4.3%	0	0.0%	0	0.0%	1	0.6%	1	0.7%
Other	195	40.2%	93	22.1%	59	37.8%	34	24.3%	75	48.7%	29	20.7%	61	34.9%	30	21.4%
Product related	140	28.9%	64	15.2%	47	30.1%	24	17.1%	37	24.0%	19	13.6%	56	32.0%	21	15.0%
All items/Articles	*485*	*100.0%*	*420*	*100.0%*	*156*	*100.0%*	*140*	*100.0%*	*154*	*100.0%*	*140*	*100.0%*	*175*	*100.0%*	*140*	*100.0%*

Panel E: External data

	Total				NYT				USAT				WSJ			
Category	Items	%	Art.	%	Items	%	Art.	%	Items	%	Art.	%	Items	%	Art.	%
Analyst	652	41.7%	326	77.6%	172	40.8%	95	67.9%	233	37.8%	119	85.0%	247	47.1%	112	80.0%
Capital Market	642	41.1%	289	68.8%	140	33.2%	82	58.6%	313	50.8%	121	86.4%	189	36.1%	86	61.4%
Competitor	165	10.6%	39	9.3%	78	18.5%	15	10.7%	33	5.4%	12	8.6%	54	10.3%	12	8.6%
Industry	88	5.6%	50	11.9%	28	6.6%	20	14.3%	29	4.7%	14	10.0%	31	5.9%	16	11.4%
Macro Data	15	1.0%	10	2.4%	4	0.9%	3	2.1%	8	1.3%	4	2.9%	3	0.6%	3	2.1%
All items/Articles	*1,562*	*100.0%*	*420*	*100.0%*	*422*	*100.0%*	*140*	*100.0%*	*616*	*100.0%*	*140*	*100.0%*	*524*	*100.0%*	*140*	*100.0%*

Notes: This table shows the number of items in the different groups (Panel A). Panels B to D split the groups into the various categories: accounting data (Panel B), and thereof the performance items (Panel C), non-accounting data (Panel D), and external data (Panel E). The analysis is based on 420 articles (Art.) from three news outlets (140 each). The percentage numbers for items refer to all items collected in a category, while the percentage number of articles refer to all sample articles (420 or 140 each, respectively).

Panel D shows the non-accounting data group. The largest category in this group refers to product-related information like the current production (28.9%). Another important category relates to employees, for which news outlets supply information like wage increases. For nearly all categories of this group, the number of items and articles does not vary substantially between the different news outlets. There are 485 items in the non-accounting data group, which account for only 6.5% of all items in the sample. This result suggests that non-accounting items only play a subordinate role in articles about earnings announcements, which might also be since firms disclose less non-accounting information when announcing their earnings. Furthermore, this category might be of less interest to the readers right after an accounting event.

The external data group (Panel E) is the second largest group with 1,562 items (21%). The highest proportion in this group belongs to analysts. Journalists mention this category in 326 (77.6%) out of 420 articles. In line with the sources of journalists state in Call et al. (2018), news outlets often use the analyst forecasts as a benchmark for the expected performance and compare them to the reported items to evaluate the performance of the firms and thus verbally state the forecast error. USAT articles contain analyst information in about 85% of the sample, while the proportion is smaller but still comparably high for the other two outlets (WSJ: 80%; NYT: 67.9%). In total, WSJ articles contain most analyst items. Additionally, journalists report the capital market reaction based on the earnings announcement or a prior long-term trend in 68.8% (298) articles. The proportion of capital market information is larger for the USAT than for the WSJ and NYT (86.4% vs. 58.6% and 61.4%). Again, this result matches the sources states in Call et al. (2018) survey paper.

2.5.5 Quotes

Table 2.9 shows the descriptive results for the content analysis of the quotes. These results are especially interesting because the quotes can either emanate from public information such as the earnings announcement or private communication, e.g., with firm executives. The former would be in line with information dissemination, while the latter is a strong indicator of information production. The results show that about 72% of all sample articles feature a quote from a firm executive. I find analyst quotes in about 25% of all articles, while the use of third-party experts is relatively rare (3%).

Table 2.9, Panel A shows the number of quotes for the internal, analyst, and external quotes in total and separately by news outlet. Though the WSJ contains most quantitative items per article, it includes fewer quotes from all categories than the other two outlets. This is surprising, as one would not expect that the access to analysts or firm executives differs for WSJ journalists. The supply of executive quotes is comparable for the NYT and USAT, but the NYT includes more analyst quotes.

About 25% of all quotes in articles emanate from analysts. As already outlined in a prior section, the WSJ includes most analyst items. However, the other outlets supply nearly twice as many quotes from analysts. Therefore, it seems like the WSJ relies on quantitative items instead of statements from analysts and produces information on a different level. Additionally, Call et al. (2018) state that some analysts do not want to be quoted in financial press articles, which might be another explanation for this diverging result.

Table 2.9, Panel B splits the results for executive and analyst quotes into the stated sources[9] and thus clearly splits information dissemination and production.[10] While prepared remarks reflect information dissemination (e.g., because they are based on the earnings announcement), the other three sources (interview, conference call, other (private) communication) in Panel B relate to information production. I further split the executive quotes into different management positions. Overall, most executive quotes derive from the conference call or prepared remarks. Nevertheless, about one-third of them are based on an interview or private communication, which is indicative of the clearest form of information production, as journalists include information that is not publicly available. The sources of the quotes differ slightly for the different executives. For example, nearly 25% of CFO quotes are based on an interview, while the CEO is more often quoted from prepared remarks. Nearly 80% of all analyst quotes derive from private communication, while about 17% reference to prepared remarks like research notes. The analyst results support the argument in Call et al. (2018), which states that journalists frequently try to include private comments from analysts in their articles to supply more relevant information.

[9] When the article does not state a source, I assume the quote is based on private communication.

[10] In contrast to section 2.6, I do not reconcile the quotes word by word but rely on the source stated in the article.

Table 2.9: Quotes in financial press articles

Panel A: Quotes by news outlet

	Total		NYT		USAT		WSJ	
Category	Items	%	Items	%	Items	%	Items	%
Internal Quote	393	72.0%	163	68.2%	123	70.3%	107	81.1%
Analyst Quote	139	25.5%	70	29.3%	45	25.7%	24	18.2%
External Quote	14	2.6%	6	2.5%	7	4.0%	1	0.8%
Total	*546*	*100.0%*	*239*	*100.0%*	*175*	*100.0%*	*132*	*100.0%*

Panel B: Quotes by source

	Total		Prepared remarks		Conference Call		Interview		other (private) communication	
Category	Items	%	Items	%	Items	%	Items	%	Items	%
Internal Quote	393	100.0%	96	24.4%	166	42.2%	42	10.7%	89	22.6%
CEO	*271*	*69.0%*	86	31.7%	*105*	*38.7%*	*20*	*7.4%*	60	22.1%
CFO	*82*	*20.9%*	4	4.9%	*44*	*53.7%*	*19*	*23.2%*	15	18.3%
Other executives	*40*	*10.2%*	6	15.0%	*17*	*42.5%*	*3*	*7.5%*	14	35.0%
Analyst Quote	139	100.0%	23	16.5%	8	5.8%	0	0.0%	108	77.7%

Notes: This table shows the quotes in financial press articles by news outlet (Panel A) and by source (Panel B). Percentage numbers in Panel A refer to the total number of quotes per news outlet (e.g., NYT). Percentage numbers in Panel B refer to the total number of items for the respective quote category (e.g., internal quotes). The analysis is based on 420 articles from three news outlets (140 each).

2.6 Information production in financial press articles

My third research question relates to the potential production role of the press. To answer this question, I perform two different analyses building on the content analysis of the previous section. First, I analyze how many of the quantitative items are based on information dissemination and production, respectively. Second, I provide evidence on the firm characteristics increasing the amount of information production and therefore estimate a determinants model.

2.6.1 Breakdown of supply into dissemination and production

As a first step, I include references on sources of quantitative items in my analysis. This way, I can infer the extent of information production (Information set II and III) and dissemination (Information set I). To that end, I reconcile the previously collected quantitative internal items with the earnings announcement press release and the quarterly conference call.[11] Earnings announcement press releases (Basu et al. (2013)) and the conference call (Matsumoto et al. (2011)) are the most important form of disclosure types for an earnings announcement. Journalists in Call et al. (2018) state they use both disclosure types as sources when writing an article. Earnings announcement press releases (Information set I) are readily available to investors, while the collection of conference call information is more time-consuming. I consider conference call information to belong to Information set II, as it does not only contain firm disclosure but also analyst opinions and information disclosed at analyst requests.

By reconciling the different items, I can establish whether the preponderance of the information derives from firm disclosure or additional sources and can clearly distinguish between information production and dissemination items. To reconcile the item, I collect the earnings announcement data from the SEC (earnings announcement press release filed as 8-K) and from Thomson Reuters Street events (conference calls transcripts). If the firm does not disclose the item in the earnings announcement, I check if the item is discussed in the quarterly conference call. If both sources do not contain the item, I assign the item to "other sources". The newspaper archives are an example for other sources, as journalists tend to include information in new articles, which has already been part of a previous

[11] I do not reconcile the quotes, as they are partly based on private communication and thus not necessarily reconcilable. I only examine sources of external items when they journalists state them directly in the article.

article about the firm. Another alternative source might be a private conversation with a CEO or similar. Overall, I assume that "other sources" mainly refers to the information set III of Figure 2.1 and is indicative of information production.

Table 2.10: Sources of internal items in financial press articles

Panel A: All items							
	N (Items)	Earnings Announcement		Conference call		Other	
		N (Items)	%	N (Items)	%	N (Items)	%
All categories	5,885	5,026	85.4%	295	5.0%	564	9.6%

Panel B: Accounting data							
Category	N (Items)	Earnings Announcement		Conference call		Other	
		N (Items)	%	N (Items)	%	N (Items)	%
Balance Sheet	33	23	69.7%	1	3.0%	9	27.3%
Cash-Flow	35	25	71.4%	7	20.0%	3	8.6%
Expenses	206	129	62.6%	29	14.1%	48	23.3%
Performance	2,033	1,962	96.5%	17	0.8%	54	2.7%
GAAP	1,617	1,571	97.2%	9	0.6%	37	2.3%
Non-GAAP	284	265	93.3%	6	2.1%	13	4.6%
EB	132	126	95.5%	2	1.5%	4	3.0%
Industry-specific	253	203	80.2%	15	5.9%	35	13.8%
M&A	65	7	10.8%	5	7.7%	53	81.5%
Revenue	1,208	1,104	91.4%	30	2.5%	74	6.1%
Segment	1,464	1,258	85.9%	77	5.3%	129	8.8%
Share related	103	65	63.1%	10	9.7%	28	27.2%
Total	*5,400*	*4,776*	*88.4%*	*191*	*3.5%*	*433*	*8.0%*

Panel C: Non-accounting data							
Category	N (Items)	Earnings Announcement		Conference call		Other	
		N (Items)	%	N (Items)	%	N (Items)	%
Employees	102	32	31.4%	18	17.6%	52	51.0%
Exchange rate	37	26	70.3%	9	24.3%	2	5.4%
Legal	11	3	27.3%	1	9.1%	7	63.6%
Other	195	114	58.5%	32	16.4%	49	25.1%
Product related	140	75	53.6%	26	18.6%	39	27.9%
Total	*485*	*250*	*51.5%*	*86*	*17.7%*	*149*	*30.7%*

Notes: This table shows the sources of the quantitative items for all items (Panel A). Panels B and C further split the internal items into groups accounting data (Panel B) and non-accounting data (Panel C). The analysis is based on 420 articles from three news outlets (140 each).

Table 2.10 shows the results for all items in Panel A and (non-) accounting data in Panel B and C, respectively. Overall, I reconcile about 90% of all items with the earnings announcement and the conference call. Most of the items (85.4%) derive from the earnings announcement, while about 5% are based on the conference call. 9.6% of the items originate from other information (Panel A). In total, I reconcile nearly 89% of all accounting items with the earnings announcement (Panel B). 97.2% of all GAAP items are reconcilable with the press release, which is the highest percentage of all categories. The portion for the other two performance categories is a bit lower but still above 93% and thus higher than any other category.

While the number of reconcilable items is highest for the performance measures, it is lowest for M&A activities (18.5%). Journalists often include previous acquisitions into the articles, though the firms do not report them again in their quarterly disclosure. M&A activities of large firms likely receive media coverage, and therefore the journalists can include their previous research in their current articles. Apart from the M&A category, the main source for all categories is the earnings announcement. For the categories cash-flow and expenses, the conference call is another relevant source to enhance the information in the articles. Surprisingly, the balance sheet category has the highest percentage of irreconcilable items (27.3%), though the absolute value of this category is rather small (N=33, N other=9).

For the non-accounting items (Panel C), the earnings announcement and the conference call are the major sources, but the number of items with other sources is much higher than for the accounting items (30.7% compared to 8%). While firms discuss exchange rate effects in their earnings announcement (70.3%) or their conference call (24.3%), the percentage of items based on firm disclosure decreases substantially for the other categories. Especially employee and legal items are based on other sources. These results are less surprising when looking at some examples of each category. The legal category includes lawsuits against the firms and settlements in fraud cases. While the employee category includes information like wage increases, it also features layoffs and executive payments, including their bonus. Particularly excessive executive pay receives negative press coverage, and the press monitors firm behavior in this category (Core et al. (2008); Kuhnen and Niessen (2012)). Overall, the two categories often contain negative firm news. While firms do not wish to draw attention to these events, the newspapers refer to them in their articles, as information production for these topics seems to be of general interest to the reader and consequently meets their demand.

Table 2.11: Sources of external items in financial press articles

Panel A: External data					
Category	N (Items)	with source	%	without source	%
Analyst	652	391	60.0%	261	40.0%
Capital Market	642	4	0.6%	638	99.4%
Competitor	165	0	0.0%	165	100.0%
Industry	88	27	30.7%	61	69.3%
Macro Data	15	0	0.0%	15	100.0%
Total	*1,562*	*422*	27.0%	*1,140*	73.0%

Panel B: Sources of external data			
Source	N (Items)	%	% cumm.
Thomson Reuters	272	64.5%	64.5%
FactSet	52	12.3%	76.8%
S&P	33	7.8%	84.6%
Zacks	20	4.7%	89.3%
IDC	10	2.4%	91.7%
Other	35	8.3%	100.0%

Panel C: Distribution of external data and sources			
Category	Analyst	Capital Market	Industry
Items	652	642	88
Sources			
Thomson Reuters	272	0	0
FactSet	52	0	0
S&P	27	4	2
Zacks	20	0	0
IDC	0	0	10
Other	20	0	15
with source	*391*	*4*	*27*

Notes: This table shows the external data items (Panel A), the sources of the external data (Panel B), and the distribution of sources and external data (Panel C). The analysis is based on 420 articles from three news outlets (140 each).

Table 2.11 (Panel A) shows the results for the sources stated for the external items. For most of the external items, the journalists do not report a source directly. Capital market information (e.g., stock prices) is the second-largest external group and easily observable. Thus, journalists might not feel the need to state a source. Industry-related information and macroeconomic data are also publicly available and, therefore, do not necessarily need a source. Journalists in Call et al. (2018) state that they compare firm performance to their industry peers. In line with their argument, I collect 165 items related to competitors. Though the outlets do not state a source in these cases, the information likely emanates from earnings announcements of the peer firms.

Journalists state items of the largest external category, analysts, with a source in nearly 52% of the cases. The most important source for analyst information is Thomson Reuters (Table 2.12, Panel B and C), as this information provider contains the aggregated analyst forecasts. Fact Set and S&P are other important sources but cited considerably less than Thomson Reuters.

Overall, the results suggest that financial press journalists use earnings press releases (Information set I) as their primary information source and rely on the information provided by the firms. This way, they act according to the information dissemination function of the financial press and rebroadcast publicly available information to a broader public. Nevertheless, some news outlets include additional sources like analyst conference calls (Information set II) and various other sources (Information set III) in their articles. This information production might facilitate information collection for investors, which might create additional value for the target group of the news outlets.

2.6.2 Determinants of information production supply

2.6.2.1 Design

In a last step, I estimate a determinants model to shed light on the characteristics of firms for which journalists, consistent with information production, choose to include additional sources like analyst conference calls, and thus enhance the information in their articles. To provide evidence on the coherence of the earnings press release and financial press article information, I first calculate the percentage of items that can I reconcile with the corresponding earnings announcement (*EA_cor*) for each article. Table 2.A4 (Appendix) shows the distribution of the mean proportion of quantitative items per article that can be reconciled with earnings press releases across the different news outlets and in total. The overall mean is 66% (NYT: 65%; USAT: 63%; WSJ: 70%).

I use *EA_cor* to calculate my dependent variable, *Add_supply,* using a median split. The median in the sample is 0.706. *Add_supply* equals 1 when the news outlet offers a higher than the median percentage (below 0.706) of information that is not based on the earnings announcement press release and 0 otherwise (above 0.706). Generally, I expect those firm characteristics that drive the coverage in financial press outlets to also influence the amount of time the journalists invest in writing their articles and, consequently, the amount and sources of article items. Therefore, I use the three sets of test variables (information environment, information content of the earnings announcement, and general firm fundamentals) outlined for the coverage regression and estimate the regression on the article level. The sample of this analysis consists of 420 observations. Table 2.A1 summarizes the variables, while Table 2.A5 (Appendix) reports descriptive statistics for this regression.

2.6.2.2 Results

Table 2.12 reports my findings from estimating a probit regression using the test variables form my three sets of variables used in the coverage analysis.[12] I estimate the probit model using quarter-year fixed effects and cluster the standard errors at the firm level. Columns (1), (5), and (6) use the additional information supply in all papers as the dependent variables, while columns (2) to (4) present the results for the three different outlets.

I find a significantly negative association between the additional information supply and the sentiment of the prior media coverage (*AES*) for all papers. This association suggests that journalists limit their information production for firms without negative news over the prior quarter. One reason for this might be that journalists consider the firms to be of less interest to their readers due to the lack of controversy. When splitting the sample into the different outlets, the results remain negative but lack a significant association with the dependent variable[13].

The coefficient for *Size* shows a positive sign and is highly statistically significant (1% level). This hints that journalists engage in further research and include additional sources when writing articles about larger firms, which suggests that information production is more pronounced for larger firms. These results are consistent for the NYT and the USAT. Coefficient estimates of *BTM* are negative and significant on a 1% level in all model specifications. I find a comparable result when estimating the model for NYT articles, but find no significant associations for the other two outlets. This indicates that NYT journalists are less likely to supply additional information about potential growth firms. Share turnover is positive and significant (1% level). This suggests that newspaper outlets include additional items when a firm is highly traded, and thus, more potential readers might be interested in quantitative items.

[12] Due to multicollinearity, I estimate the probit models without the loss firm indicator variable for the separate outlets.
[13] The samples used for the three different outlets are rather small and might therefore lack statistical power.

Table 2.12: Determinants of additional information supply

	(1) Add_supply Full sample	(2) Add_supply NYT	(3) Add_supply USAT	(4) Add_supply WSJ	(5) Add_supply Full sample	(6) Add_supply Full sample
Size	0.298***	0.305**	0.487***	0.238	0.168**	0.162**
	(3.85)	(2.03)	(2.62)	(1.33)	(2.12)	(2.02)
Freefloat	-0.318	-2.093	0.103	1.012	0.005	-0.034
	(-0.56)	(-1.39)	(0.11)	(0.68)	(0.01)	(-0.06)
BTM	-0.808***	-1.529***	-0.599	-0.202	-0.638**	-0.655**
	(-2.87)	(-2.76)	(-1.46)	(-0.41)	(-2.05)	(-2.06)
AF	0.285	0.284	0.238	0.527	0.461*	0.459*
	(1.17)	(0.56)	(0.48)	(1.25)	(1.78)	(1.76)
AFE	-4.186	-8.721	-3.524	-2.021	1.000	-0.725
	(-1.11)	(-1.41)	(-0.52)	(-0.31)	(0.21)	(-0.15)
SVI_sd	0.032**	0.037	0.027	0.024		
	(2.22)	(1.30)	(0.92)	(0.89)		
Turnover	0.061***	0.079***	0.148***	-0.034		
	(3.26)	(2.75)	(3.34)	(-0.89)		
Cov Q-1	0.009	-0.009	-0.255	-0.126	0.027	0.016
	(0.04)	(-0.03)	(-0.89)	(-0.42)	(0.11)	(0.06)
AES	-0.810*	-1.176	-1.547	-0.043	-1.227***	-1.117**
	(-1.76)	(-1.32)	(-1.47)	(-0.05)	(-2.68)	(-2.35)
BusyDay	0.155	0.199	0.176	0.184	0.125	0.110
	(0.94)	(0.61)	(0.57)	(0.66)	(0.75)	(0.66)
SVI_mean					0.124	0.131
					(0.89)	(0.94)
AF_miss					-0.659	
					(-0.77)	
Loss firm						0.409
						(0.59)
Volatility						0.051
						(0.08)
BHR						20.240
						(1.36)
Pseudo R^2	0.104	0.213	0.211	0.135	0.084	0.083
N	420	140	140	140	420	420
Fixed Effect	Quarter	Quarter	Quarter	Quarter	Quarter	Quarter

Notes: This table shows the results of a probit regression with the median split of information supply (*Add_supply*) as the dependent variable for either all papers or for the three different outlets outlined in the column's headline. All variables are defined in Table 2.A1 (Appendix). Variables comprise Size, Freefloat, BTM (Book-to-market), AF (number of analyst forecasts), AFE (analyst forecast error), SVI_sd (Google search standard deviation), Turnover, prior coverage (outlets combined and separately), AES (press sentiment score), BusyDay dummy, loss firm dummy, volatility, AF_miss (dummy for missed forecast), and BHR (buy-and-hold return). ***, **, * indicate statistical significance at the 1%, 5%, and 10% level, respectively, using two-tailed tests and standard errors clustered at the firm level. (N articles=420). Maximal VIF in all models is 2.68.

In contrast to the coverage analysis, the coefficients of *Cov Q-1* are not significant. This suggests that though journalists keep covering a firm, it does not render it more likely that they produce additional value beyond the dissemination of information. The standard deviation of the Google search index (*SVI_sd*) has a positive association with additional information. This suggests that news outlets partly anticipate the demand of their readers and provide more information when interest varies in the prior quarter, whereas the mean interest in the firm yields no significant results (*SVI_mean*). However, coefficient estimates of *SVI_sd* yield no significant results on the outlet level.

Overall, the firm characteristics determining additional effort of the journalists seem similar to the determinants of the general coverage and mostly relate to the general information environment of a firm. Additional analyses indicate that less specialized newspapers devote more resources to large firms with more liquidity, as their readers might be more interested in well-known firms, whereas the instances in which the WSJ produces additional information are less predictable.

2.7 Conclusion

Prior literature supports the role of the financial press as an information intermediary, whose coverage increases trading volume and reduces information asymmetry (e.g., Bushee et al. (2010); Blankespoor et al. (2018); Coyne et al. (2019)). However, the evidence about how the press creates value is mixed. While some papers limit the value to information dissemination, others argue that the press also produces valuable information. My findings on the supply of information by the financial press contribute to a growing yet conflicting literature on the economic role of the financial press. First, to my knowledge, I am the first to provide detailed, content analysis-based evidence on the choice of quantitative information items to be reported. The comparison of a specialized newspaper and two general-interest newspapers provide novel insights into the supply of intermediation activities to the market, suggesting that the newspapers cater to their different audiences.

Second, and related, my analysis of press article content enables me to discern information dissemination from information production in the financial press and to assess how the press fulfills each role. I show that the dissemination function is shaped by demand, in particular for information on large firms and primarily spreads earnings and revenue measures. Notably, though, my findings also suggest that there is a demand for additional information provided by the press, by assuming a production function, e.g., based on additional information on M&A activities or expenses. My results imply that information production is not limited to quantitative information, as the quotes feature further insights on the verbal level.

In this paper, I present explorative evidence on the supply of information intermediation provided by the financial press. Based on the analysis of three different news outlets that cover quarterly earnings announcements, I show that the supply and width of information by the financial press on accounting events are associated with the richness of the respective firm's information environment.

When breaking down the content of the articles based on the sources of information items included, my analyses reveal that the financial press assumes its dissemination function by channeling information from earnings announcements to market participants. More importantly, I show that the press also assumes an information production role by combining information from various other sources, e.g., conference calls. The inclusion of analyst and executives quotes from interviews expand the information production from quantitative items to the verbal level.

One limitation of my analyses is that I can draw no inferences about what kind of information newspapers might produce beyond the quantitative items. As my results for the quotes indicate, journalists frequently include information on a verbal level without reference to "hard information". The quotes are regularly based on interviews or conference calls and thus produce information. Consequently, there is more research needed on factors that shape the demand for additional (verbal) content provided in press articles and how the supply of this content is economically useful to readers of financial newspapers. That said, I need to caution that my findings are based on a comparatively small sample of generally large firms and large news outlets with potentially highly qualified journalists. Furthermore, I do not provide evidence about what kind of information news outlets decide not the include in their articles, though firms presented items or verbal explanations about the topics in their disclosure. Also, given the descriptive nature of my analyses, I do not claim to provide a comprehensive explanation of the factors that shape demand for financial press intermediation in a causal sense.

2.8 Appendix

Appendix 2.A1: Variable definition

Variable Name	Definition	Source
Add_supply	Indicator variable set to one if EA_cor is above the median and zero otherwise	Manual collection
AES	The ratio of positive events reported about a firm compared to the sum of positive and negative events within a 91-day window (set to neutral if missing)	Ravenpack
AFE*	Analyst Forecast Error using the median forecast, scaled by the stock price at the quarter-end	IBES
BHR	Buy & Hold return for the previous quarter beginning on the day of the previous earnings announcement and ending one day before the current earnings	CRSP
BTM	Book-to-market ratio	Compustat
BusyDay	Indicator variable set to one if the number of earnings announcements per day is above the median and zero otherwise	Compustat
Coverage	Indicator variable set to one if the press covers the firm in the respective quarter (further split into the different newspapers)	Ravenpack
Cov Q-1	Indicator variable set to one if the earnings announcement has been covered in the previous quarter (further split into the different newspapers in subsample analyses)	Ravenpack
EA_cor	Percentage of items that can be reconciled with the earnings announcement (except quotes)	Manual collection
Freefloat	Free float of total shares outstanding	Datastream
SVI_mean	Natural logarithm of one plus the SVI mean of the four weeks prior to the earnings announcement	Google
SVI_sd	SVI standard deviation of the four weeks prior to the earnings announcement	Google
Loss_firm	Indicator variable set to one if the firm reports a loss in the current quarter	IBES
AF_miss	Indicator variable set to one if the firm missed the analyst forecast for the current quarter	IBES
AF	Natural logarithm of one plus the number of analyst EPS estimates	IBES
Size*	Natural logarithm of market capitalization	Compustat
Turnover	Mean share turnover during the previous quarter beginning on the day of the previous earnings announcement and ending one day before the current earnings announcement	CRSP
Volatility	Return volatility during the previous quarter beginning on the day of the previous earnings announcement and ending one day before the current earnings announcement	CRSP

Notes: This table shows variable definitions. * denotes the variable is winsorized at the 1% and 99% level due to outliers.

Appendix 2.A2: Correlation analysis

Panel A: Regression variables

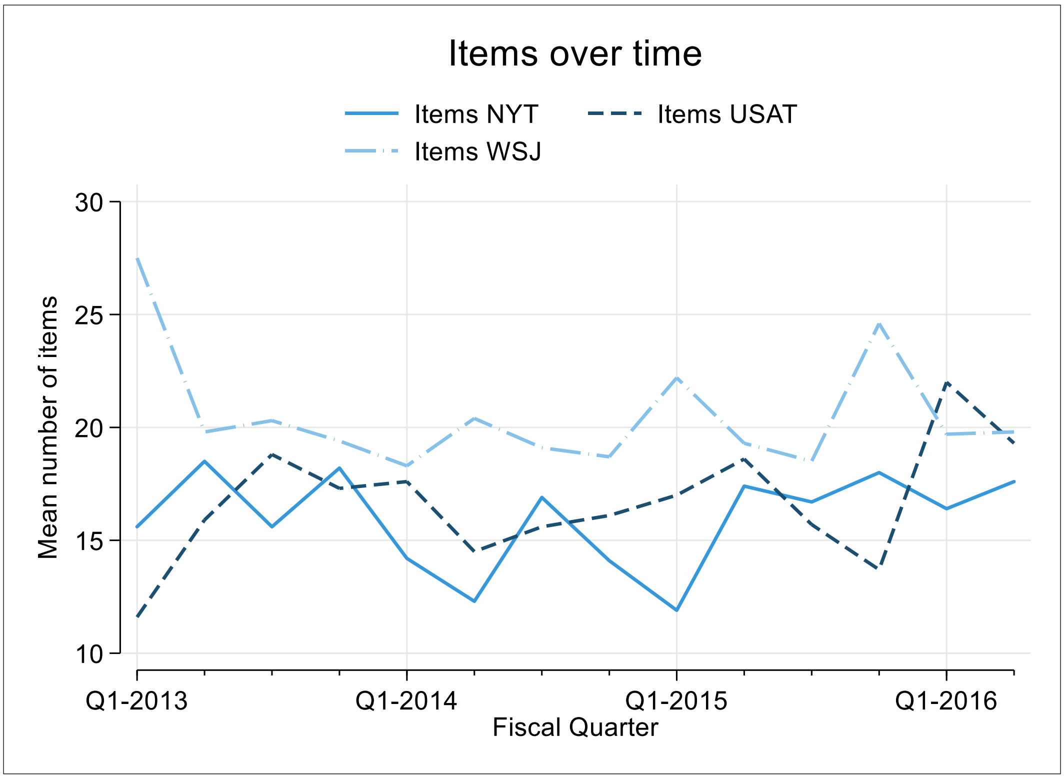
	Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	Size	1.00														
2	Freefloat	0.18	1.00													
3	BTM	-0.09	0.11	1.00												
4	AF	0.33	0.05	-0.10	1.00											
5	AFE	0.01	-0.16	-0.38	-0.04	1.00										
6	AF_miss	-0.08	0.00	0.11	0.03	0.37	1.00									
7	AES	0.22	0.04	-0.08	0.12	-0.02	-0.11	1.00								
8	loss_firm	-0.07	-0.03	0.10	0.03	0.24	0.53	-0.13	1.00							
9	SVI_mean	-0.01	0.08	0.04	-0.10	-0.07	0.01	0.01	-0.01	1.00						
10	BusyDay	-0.06	0.03	-0.02	-0.06	0.01	0.00	-0.01	-0.01	0.03	1.00					
11	Turnover	-0.41	-0.10	0.12	0.10	0.00	0.23	-0.17	0.25	-0.04	-0.03	1.00				
12	SVI_sd	-0.20	-0.02	0.01	-0.11	0.03	0.02	-0.06	-0.01	0.10	0.01	0.10	1.00			
13	Volatility	-0.24	-0.12	0.11	0.09	0.04	0.28	-0.16	0.31	-0.07	-0.02	0.70	0.10	1.00		
14	BHR	0.01	-0.02	-0.06	0.01	0.05	0.01	0.09	0.01	0.00	0.04	0.00	0.00	-0.06	1.00	
15	Coverage Q-1	0.41	0.05	-0.08	0.27	-0.05	-0.03	0.23	-0.01	0.03	-0.11	-0.02	-0.09	0.05	0.00	1.00

Panel B: News outlets

	Variable	1	2	3	4
1	coverage	1.000			
2	wsj_cov	0.968	1.000		
3	usat_cov	0.368	0.306	1.000	
4	nyt_cov	0.320	0.279	0.588	1.000

Notes: This table shows the variable correlation for the coverage analysis (Panel A) and the newspaper coverage (Panel B).

Appendix 2.A3: Quantitative items over time



This figure shows the mean number of quantitative items overall sample quarters for the random sample. I analyze 10 articles each quarter for the three news outlets (N=30 per quarter, N=140 per news outlet, N=420 in total) between Q1-2013 and Q2-2016.

Appendix 2.A4: Coherence of press articles and earnings announcements

	Total		NYT	USAT	WSJ
Quarter	SD	Mean	Mean	Mean	Mean
2013q1	0.2	0.66	0.62	0.7	0.66
2013q2	0.15	0.73	0.71	0.72	0.75
2013q3	0.2	0.65	0.64	0.65	0.65
2013q4	0.2	0.66	0.58	0.68	0.72
2014q1	0.17	0.64	0.61	0.55	0.76
2014q2	0.23	0.65	0.64	0.61	0.68
2014q3	0.22	0.66	0.63	0.63	0.73
2014q4	0.23	0.62	0.66	0.54	0.67
2015q1	0.18	0.71	0.71	0.66	0.75
2015q2	0.18	0.66	0.68	0.6	0.68
2015q3	0.15	0.69	0.67	0.65	0.74
2015q4	0.2	0.62	0.69	0.5	0.66
2016q1	0.18	0.63	0.62	0.64	0.63
2016q2	0.16	0.7	0.69	0.72	0.69
Total	*0.19*	*0.66*	*0.65*	*0.63*	*0.7*

Notes: This table shows the mean coherence (*EA_cor*) between earnings announcements and financial press articles during the sample period for the three sample outlets (N=420, N=140 per news outlet).

Appendix 2.A5: Random sample and summary statistics (content analysis)

Panel A: Sample distribution per industry (for the random sample)

Fama-French industry code (12 industries)	Firms	Firm-quarters	Percent	Percent (all firms)
(1) Consumer Non-Durables	29	336	6.9%	7.4%
(2) Consumer Durables	7	98	1.7%	2.2%
(3) Manufacturing	31	336	7.4%	7.4%
(4) Oil, Gas, and Coal Extraction (Energy)	18	252	4.3%	5.6%
(5) Chemicals and Allied Products	11	196	2.6%	4.3%
(6) Business Equipment	93	686	22.1%	15.2%
(7) Telephone and Television Transmission	37	140	8.8%	3.1%
(8) Utilities	2	336	0.5%	7.4%
(9) Wholesale, Retail, and Some Services	80	448	19.0%	9.9%
(10) Healthcare, Medical Equipment, and Drugs	17	350	4.0%	7.7%
(11) Finance	65	812	15.5%	18.0%
(12) Other (e.g., Hotels, Entertainment)	30	532	7.1%	11.8%
Total	*420*	*457.8*	100.0%	100.0%

Panel B: Summary statistics (for the random sample)

Variable	mean	sd	p50	min	max	Count
Coverage All Papers	0.99	0.12	1.00	0.00	1.00	420
WSJ Coverage	0.97	0.17	1.00	0.00	1.00	420
USAT Coverage	0.67	0.47	1.00	0.00	1.00	420
NYT Coverage	0.63	0.48	1.00	0.00	1.00	420
Cov Q-1	0.92	0.28	1.00	0.00	1.00	420
NYT Cov Q-1	0.56	0.50	1.00	0.00	1.00	420
USAT Cov Q-1	0.55	0.50	1.00	0.00	1.00	420
WSJ Cov Q-1	0.87	0.33	1.00	0.00	1.00	420
cor_8k_all	0.67	0.19	0.71	0.00	1.00	420
AES	0.69	0.15	0.71	0.15	1.00	420
BusyDay	0.34	0.47	0.00	0.00	1.00	420
Size	11.09	1.09	11.17	8.49	12.68	420
Freefloat	0.92	0.11	0.94	0.48	1.00	420
BTM	0.38	0.30	0.30	-0.12	1.73	420
Loss firm	0.02	0.13	0.00	0.00	1.00	420
SVI_mean	3.77	0.60	3.86	0.00	4.59	420
SVI_sd	5.00	4.72	3.37	0.00	29.10	420
Turnover	7.49	5.22	6.30	0.01	43.25	420
Volatility_quarter	0.01	0.01	0.01	0.01	0.06	420
AF	3.25	0.34	3.30	1.39	4.04	420
AFE	-0.05	0.02	-0.05	-0.14	0.02	420
AF_miss	0.01	0.11	0.00	0.00	1.00	420
BHR	0.03	0.11	0.03	-0.67	0.33	420

Notes: This table provides information about the industry distribution (Panel A) and summary statistics for the variables of the coverage analysis (Panel B). Sample selection is defined in Table 2.2. All variables are defined in Table 2.A1 (Appendix). Variables comprise Coverage (outlets combined and separately), prior coverage (outlets combined and separately), AES (press sentiment score), BusyDay dummy, Size, Freefloat, BTM (Book-to-market), loss firm dummy, SVI_mean (Google search mean, SVI_sd (Google search standard deviation), Turnover, volatility, AF (number of analyst forecasts), AFE (analyst forecast error), AF_miss (dummy for missed forecast), and BHR (buy-and-hold return).

3 Information production by the financial press: A closer look

Ann-Kristin Großkopf, Jörg-Markus Hitz[14], and Harm Schütt[15,16]

Working Paper[17]

Abstract: We use topic modeling (LDA) to investigate in detail how the financial press produces information in articles that cover firms' quarterly earnings announcements. We find evidence that journalists potentially create information in a variety of ways: by adding additional topics to their coverage and by omitting presumably less important content from earnings announcements, by emphasizing specific topics, and by "toning down" firms' narratives. Market tests demonstrate that the magnitude of investor reactions to earnings news contained in earnings announcements increases with our topic-based measures of information production by the press. These findings are more pronounced for information production by way of modifying content (adding or removing topics) than for toning down firms' earnings news. Taken together, these findings suggest that information production by the press creates value to investors by helping to understand earnings news. Specifically, our findings augment prior literature by documenting in detail how journalists produce relevant information.

JEL Classification: M40, M41, G14

Keywords: Financial press, information intermediation, earnings announcement

Acknowledgments: For valuable research assistance, we thank Jana Wagner.

[14] Georg-August Universität Göttingen, Faculty of Economic Sciences, Chair of Accounting and Auditing, Platz der Göttinger Sieben 3, 37073 Göttingen, Germany.

[15] Tilburg University, Tilburg School of Economics and Management, Department Accountancy, 5037 AB Tilburg, Netherlands.

[16] This study was conducted with Jörg-Markus Hitz and Harm Schütt.

[17] First version of the working paper: February 2020. This version: March 2020.

3.1 Introduction

This study investigates how journalists produce information for their audiences, distinguishing the production of original content from disseminating firm disclosures. The financial press is an important information intermediary that potentially reaches the largest and most diverse audience among investors. A growing literature investigates this role of the financial press as an intermediary of firm information in capital markets, identifying two principal functions which the financial press assumes to create value for investors: a dissemination role and a production role. The press assumes a dissemination role by disseminating firm information, e.g., from earnings announcements, in capital markets. In comparison, the press assumes an information production role by creating or adding additional content to their articles, e.g., in the shape of analyses, or by combining information from additional sources such as analyst reports, or interviews with firm executives (Miller and Skinner (2015)). Both functions are usually intertwined within an article and can both affect how investors react to the news (Drake et al. (2014)).

There is consistent and conclusive evidence in support of a dissemination role of the financial press. Various papers investigate this link between information dissemination, investor attention, and market outcomes. These papers find that media coverage drives trading volume and is significantly related to liquidity, and costs of capital (Peress (2014); Fang et al. (2014); Engelberg and Parsons (2011); Kothari et al. (2009); Blankespoor et al. (2018)). Hence, the financial press contributes to price discovery in financial markets.

In contrast to the dissemination role, research on the information production role of the financial press is less plentiful and, more importantly, less consistent (Miller (2006); Ahern and Sosyura (2015)). Most papers (e.g., Drake et al. (2014)) have been unable to demonstrate that press article content, such as in-depth company or industry analyses, or criticizing firms' use of pro forma earnings, creates value for investors. This suggests that journalists, by supplementing firm information with additional content, may solely assume an "entertainment role" (Guest (2018)). Only very recently, two papers have emerged in support of an information production role. First, Call et al. (2018), in a survey among 462 financial journalists, document that journalists expend significant effort into sourcing additional content that they factor into their coverage, e.g., of earnings announcements. Journalists do so by informal communication (background talks, phone calls) with firm insiders (investor relations officers, executives) to glean private background information, and by interacting with financial analysts. Accordingly, journalists explicitly state that producing "exclusive content" is important to their performance evaluation by superiors. Second, the paper by Guest (2018) exploits plausibly exogenous shocks to the coverage of earnings announcements by the Wall Street Journal (WSJ). Guest (2018) provides empirical evidence of a positive, plausibly causal relation between the proportion of original analyses in press articles and price discovery, as measured, e.g. by the magnitude of earnings response coefficients.

While the recent literature helps demonstrate that journalists deliberately assume an information production role, and that this role creates value in the shape of information content, these studies are unable to answer *how* journalists assume this production role. For instance, Call et al. (2018) single out topics that journalists find particularly relevant to their audiences, such as corporate fraud or insider trading. However, Call et al. (2018) do not give a detailed descriptive account of actual content produced by journalists. Also, said topics relate to specific firm events and issues and do not explain why and how coverage of earnings announcements by journalists produces content, as demonstrated by Guest (2018). The analyses of Guest (2018), likewise, do not speak to this question either, as his measure of information production relies on the textual similarity between a firm's earnings announcement press release and the associated WSJ article. Therefore, the Guest (2018) approach is unable to discern different sources of information production, e.g., in-depth analysis of firm fundamentals versus adding other sources such as interviews or analyst reports. Also, the textual similarity measure does not discern the production of original thematic content by journalists from rephrasing or "toning down" of information originated by firms. Hence Guest (2018) is unable to link specific content (topics that journalists elect to report on) to capital market outcomes.

Taken together, although recent papers demonstrate that journalists assume an information production role to the benefit of investors, the nature of the editorial content that journalists produce is much of a black box, and so is the market digest of this content. Our paper sheds light on this black box of information production by the financial press. Our objective is to provide a more nuanced perspective on the nature and amount of editorial content that journalists produce and to investigate how different ways of information production provide value to investors. Information production is a multi-attribute concept because there are multiple ways of producing information. For example, journalists can produce information by preprocessing or removing unimportant information, by garnering additional information via communication with firm insiders and experts, or by conducting their own information gathering and analysis. We use topic analysis to measure four sub-activities related to information production: 1) adding content (adding topics to articles that were not contained in the earnings announcement), 2) omitting content (omitting topics in articles that were contained in the earnings announcement), 3) emphasizing important topics (assigning more relative space to specific topics in the press article, compared to the earnings announcement), and 4) applying sentiment to topics in financial press articles. These measures enable us to address two intertwined research questions: How do journalists produce information by creating editorial content, and how do these production activities help investors analyze firms' earnings news?

We examine a sample of earnings announcements of S&P 500 firms between 2010 and 2016 and corresponding WSJ articles. We identify topics, i.e., narratives of related thematic content, using a Bayesian topic-modeling algorithm termed Latent Dirichlet Allocation (LDA). LDA enables us to efficiently extract topics from a large sample of earnings announcements and corresponding press articles. Further, the application of LDA reduces researcher bias: Unlike content analysis techniques, it does not require subjective classification decisions on behalf of the researchers. Also, as a "self-learning" algorithm, LDA does not rely on a pre-determined, potentially biased dictionary of words (Brown et al. (2020)). Based on 35 unique topics, we apply the LDA model to our sample of 6,540 press articles and earnings announcements pairs. We further aggregate these 35 topics by related content to arrive at 18 unique topic groups.

Armed with our topic-based measures of information production, we first explore the magnitude of the four dimensions of information production in the press articles to provide evidence on how journalists attempt to create value. Our first two measures examine how journalists produce information by adding or by omitting topics in the press articles compared to the corresponding earnings announcement. Earnings announcements in our sample, on average, cover 9.4 topics. We find that, on average, journalists omit 5.5 of these topics in their corresponding press articles. While space constraints in press articles also necessitate such topic reduction, it reflects at the same time, the objective of focusing on topics relevant to the investor audience. In comparison, and again consistent with the focused nature of journalists' narratives, press articles include an average of 1.5 additional topics. These additional topics reflect an important aspect of information production, as including supplementary content is a strong signal to readers about the importance of these specific topics. Our analyses demonstrate that the topic that journalists add to articles most frequently are statements from firm executives or other experts such as analysts, which underscores the value of background information to the target group, i.e., the readership of the WSJ.

To examine our third proxy of information production, journalists (de-)emphasizing specific topics in their articles, we focus on joint topics, i.e., topics that are featured both in the press article and in the earnings announcement. We measure emphasis by examining the relative proportion that journalists devote to one specific topic, compared to the earnings announcement. This analysis unveils a relatively high degree of variation in emphasis added by journalists between topics, which illustrates that journalists make extensive use of quantitative emphasis as a source of information production. For example, while journalists on average emphasize topics such as “Sales” or “BusinessOutlook”, they devote less article space to topics like “Risk” or “Segments”.[18]

In contrast to the first three measures of information production that capture various dimensions of content added, our fourth measure, sentiment, captures *how* this information is presented. Consistent with a generally skeptical approach to firm information, we find that, on average, press articles are significantly more negative than the corresponding earnings announcements. When splitting the article sentiment for added and joint topics, we find that joint topics are, on average, communicated with a more negative tone compared to added topics. However, the variation in tone is also larger for added topics, suggesting that the sentiment of added topics is more contextual compared to topics taken from the earnings announcement. Taken together, our first set of analyses documents that journalists use various article attributes to produce information for their readers. This underlines the need to explore in more detail how markets respond to these different channels of information production.

Building on our descriptive results, our second set of analyses investigates whether and with what magnitude the different forms of information production are associated with market reactions, holding information content fixed. For these multivariate analyses, we aggregate our production measures on the article level and explore three of our production measures: added content, omitted content, and sentiment. For the first measure, we find

[18] The topic „Sales“ features revenue related information, while “BusinessOutlook” includes firm guidance, e.g., earnings forecasts. The “Risk” topic comprises information about factors impacting the future performance, while “Segments” features content about the operating segments of the firms.

that adding original content to the articles is associated with a stronger investor response to earnings announcements. This association increases with the proportion of additional content contained in the articles. We obtain very similar results for information production via our second measure, the omission of topics. Market reactions are larger for earnings announcements covered by articles that omit topics, and this effect increases with the focus of the journalists, i.e., with the amount of content that journalists decide not to report.

Third and last, we analyze the impact of information production through the sentiment of an article and find that the earnings response increases with journalists' sentiment. However, when we split the sentiment contained in press articles by distinguishing tone for additional topics from tone for jointly covered topics, we find that the positive association with the earnings response is confined to joint topics only, i.e., topics that pick up content from the earnings announcement. This result suggests that the "toning down" of earnings announcement topics produces more value relevant content compared to how journalists present their additional information. Overall, our results imply that journalists use various financial press article attributes to produce information, which alters how investors respond to the earnings news. However, the nature of our tests does not allow for causal inferences, as the decision (not) to cover an earnings announcement of the firm is endogenous.

Our paper contributes to the literature on the role of the financial press as an intermediary of accounting information. Going beyond Guest (2018), our topic modeling approach enables us to measure how journalists produce information in their articles, and how this creates value for their target audience. Our findings suggest that information production in financial press articles is a multi-attribute concept and that both principal production avenues, modifying content by adding or omitting topics, and toning down firms' narratives, facilitate the interpretation of reported earnings for investors, albeit at economically different magnitudes. While our findings cannot rule out that the financial press also assumes an entertainment function, we find consistent evidence that the various editorial modifications that we examine all create value to investors.

The paper proceeds as follows. In Section 3.2, we outline our topic-based measurement of information production. Section 3.3 presents our descriptive analyses of information production by the financial press, and Section 3.4 presents results from market tests. Section 3.5 concludes.

3.2 Topic-based Measurement of Information Production

3.2.1 Measurement of Information Production

Graf-Vlachy et al. (2019) argue that press articles have various attributes such as volume, tone, topic, and other language characteristics. Consequently, there are multiple ways for the press to produce information and information production thus is a multi-faceted concept. For example, journalists can produce information by preprocessing or removing unimportant information, by garnering additional information via communication with firm insiders and experts, by conducting their own information gathering and analysis, or by applying a distinctly different tone of language than the one adopted in firm publications. Using topic analysis enables us to identify four such different activities related to information production, hence creating four unique measures.

Our first measure is *additional content* (*Add_Cont*). This measure identifies articles that journalists augment by adding additional content not included in the corresponding earnings announcements. By adding topics, journalists effectively produce additional information, since they "add" a signal about the relevance of certain topics versus others. For example, journalists may choose to supplement the discussion of a firm's financials with comparative data for industry peers, taken from other sources (e.g., databases). As another example, journalists may add an analysis of governance-related problems, e.g., a critique of inefficient director monitoring based on background analyses to their discussion of a firm's underperformance. In Call et al. (2018), 75% of the journalists state providing exclusive content impacts their evaluation by their superior. Consequently, they have an incentive to offer background information, such as analyst opinions. On the topic level, we measure *ADD_CONT* as an indicator variable, which equals one if the topic is featured in the press article but not in the earnings announcement and zero otherwise.

Our second measure is *omitted content* (*Omit_Cont*). We measure irrelevant content based on whether journalists decide not to cover topics in their article that are included in the corresponding earnings announcement. The motivation behind *Omit_Cont* is analogous to the motivation behind *Add_Cont*: By choosing which topics to neglect, journalists effectively produce additional information, since they "add" a signal about the (ir)relevance of certain topics versus others. In that sense, the interpretation of *Omit_Cont* is twofold: Journalists signal that particular topics covered in the earnings announcements are not relevant for the informational needs of readers, and / or journalists want to devote more of the scarce space available for their article to topics that they deem particularly important. As firms tend to extend their disclosure over time (e.g., Dyer et al. (2017)), focusing the article on important information might prevent the distraction of investors. Comparable to *Add_Cont*, we measure *Omit_Cont* as an indicator variable, which equals one if the topic is included in the earnings announcement but not in the corresponding press article.

Our third measure is the emphasis on specific topics (*Emphasis*). We define *Emphasis* as the topic proportions in articles vs. earnings announcements. This enables us to assess in detail which information is deemed important enough to allocate a bigger portion of the article to it. For example, the additional analysis could result in a higher topic proportion in the article versus the earnings announcement. We expect the press to engage in information production in the form of a stronger emphasis in cases where journalists judge a few pieces of information disclosed to be significantly more important than the remaining earnings announcement disclosure. This emphasis creates value relevant information either by signaling which information is important and which is not, by providing a detailed discussion of the topic or by explicitly providing additional analysis. All three aspects would lead to additional information being produced. However, as *Emphasis* relies on joint topics, i.e., topics featured in both the press article and earnings announcement, this proxy partly captures information dissemination through financial press articles. We measure *Emphasis* as the difference between the topic proportion of a specific topic in the press article and the proportion for the same topic in the earnings announcement.

Our first three measures *Add_Cont*, *Omit_Cont*, and *Emphasis* are purely content-based and measure to what extent journalists produce information by adding, reducing, or focusing on topics in their articles. Therefore, in essence, these measures capture *what* is being disclosed by the financial press. In contrast, our fourth and final measure, *Sentiment*, analyzes *how* journalists choose to cover specific topics (Brown et al. (2020)). The information content of financial press articles is not limited to the topics covered in the article but includes the way the journalists present the information, e.g., the readability or the sentiment conveyed in the article (Graf-Vlachy et al. (2019)). Huang et al. (2014) show that managers can use the tone of earnings announcements to convey biased signals to the market. Goldman et al. (2019) argue that it is the task of journalists to “debias the announcement” to limit the exposure of readers to biased information. To that end, Tetlock (2007) and Tetlock et al. (2008) provide evidence that negative financial press articles convey information about negative capital market developments and lower subsequent firm performance beyond analyst recommendations or firm fundamentals. Consistent with this, Kothari et al. (2009) suggest that negative press coverage is associated with higher cost of capital. Therefore, the *sentiment* of financial press articles is another channel for journalists to produce information.

We measure the sentiment of articles (*Sent_Art*) and earnings announcements (*Sent_EA*) as the difference between positive and negative words, scaled by the sum of positive and negative words based on the Loughran and McDonald (2011) lists. For our empirical analyses, we further disaggregate *Sent_Art* into *Sent_Add_Topics*, which captures the sentiment of added content (topics that are only featured in the press article and not in the earnings announcement) and *Sent_Joint_Topics* capturing the sentiment of joint content (topics that are featured in both article and earnings announcement).

3.2.2 Topic Modeling using Latent Dirichlet Allocation

Our empirical measures that capture various dimensions of information production by the press derive from the identification and distribution of topics in press articles and earnings announcements, respectively. To obtain these topics, we use Latent Dirichlet Allocation (LDA) (Blei et al. (2003)). LDA is an unsupervised Bayesian topic modeling algorithm. It is commonly used by internet search engines to improve the search term output (Dyer et al. (2017)). Though LDA is comparably new to the accounting literature (see Eickhoff and Neuss (2017) for a review of topic models in managerial disciplines), prior literature such as Hoberg and Maksimovic (2014), Brown et al. (2020), and Dyer et al. (2017) uses this approach to obtain topics from 10-K narratives.

The LDA model has three main assumptions, which offer key advantages with regard to our research question. First, LDA is a "bag of words" method, i.e., the model assumes that the order of words in a document and the order of the documents in the sample do not matter. In contrast to a costly manual collection, LDA can be applied to large amounts of documents. Second, LDA assumes that each document is a collection of topics from a finite number of topics. The researcher defines this finite number of topics as an input factor for the LDA, e.g., based on the perplexity score of the topic model.[19] Based on this topic selection, not each document needs to contain every topic. We obtain the proportion of each topic in each document of our corpus (sample documents) from the LDA. By comparing the topics of the press articles and earnings announcements, we can distinguish between joint and added topics (Huang et al. (2017)). Third, the word distribution of each topic follows a Dirichlet distribution. For each word of the corpus, the LDA estimates its probability weight within a topic indicating its importance for this topic. A word may be associated with multiple topics. Using all words of the corpus, the LDA, therefore, does not require predetermined dictionaries or categories, which allows the algorithm to "discover" the topics independent of the researchers' expectations and to fully reflect the information conveyed in press articles and earnings announcements.

Applying the LDA to a sample of earnings announcements and corresponding press articles allows us to identify topics used in these respective publications, and from this delineate our first two measures of information production for added content (*Add_Cont*) and omitted content (*Omit_Cont*). Our third measure, *Emphasis*, examines the proportions of each joint topic, i.e., the topics that are featured in the press article and the earnings announcement. For our fourth measure, *Sentiment*, we need to delve deeper: first, by measuring the tone of each earnings announcement and corresponding press article as outlined in section 2.1, and then by attributing the tone to the narratives that cover one specific topic. To attribute the tone to a specific topic, we examine the articles on the sentence level. To that end, we use the term-topic matrix of the LDA model, which features the probability with which a word belongs to a specific topic. We use this matrix to assign each sentence in an article to a topic. To do so, we sum up the topic probabilities of each word in a sentence and allocate the sentence to the topic with the highest probability (Huang et al. (2017)). We then measure the tone of each sentence using the positive and negative word lists of Loughran and McDonald (2011). We classify a sentence as positive (negative) if the number of positive words is larger (smaller) than the number of negative words. We aggregate sentences according to the coverage of the respective topic into added content and joint content sentences and measure the tone separately for each category. We measure *Sent_Add_Topics* as the difference between positive and negative

[19] Blei et al. (2003) define perplexity as $perplexity(D_{test}) = exp\left\{\frac{\sum_{d=1}^{M} log\ p(w_d)}{\sum_{d=1}^{M} N_d}\right\}$. The perplexity score evaluates the statistical fit of the overall model based on different numbers of topics by splitting all sample documents into a training data and a testing data set. A lower perplexity indicates that the model has better generalization performance.

added topic sentences scaled by all added topic sentences in an article. Similarly, we measure *Sent_Joint_Topics* as the difference between positive and negative sentences scaled by all non-production sentences in an article.[20]

3.3 Information Production by the Financial Press

3.3.1 Sample Selection

We examine quarterly earnings announcements press releases and the corresponding WSJ articles between 2010 and 2016. The WSJ is the main financial newspaper in the U.S., and it has been used in several prior studies investigating the role of the financial press as an information intermediary (e.g., Dougal et al. (2012); Tetlock (2007); Drake et al. (2014)). Following prior literature, we confine our analyses to S&P 500 firms (Tetlock et al. (2008); Guest (2018)). As these firms make up about 80% of the U.S. market capitalization, we assume they receive the necessary coverage for our analysis. We exclude all firms with missing earnings announcement dates from the sample. We identify all WSJ articles related to our sample firms around the quarterly earnings announcement day [-7/+7] using Ravenpack (Drake et al. (2014)). We exclude all articles with a Ravenpack relevance score below 70.[21] Additionally, we exclude all articles that do not feature any information about the earnings announcement, feature earnings information about multiple firms, or focus on specific non-earnings-related topics. We drop firms without WSJ coverage and missing data from the sample and end up with 498 sample firms.

We identify 7,333 WSJ articles related to our sample firms. We obtain the text of these articles from ProQuest and the earnings announcement press releases from the 8-K SEC filings. We drop articles from the sample when we are unable to obtain the corresponding press release from the SEC. Whenever there are multiple articles about an earnings announcement, we combine these articles into one article observation. We collect firm fundamentals like the earnings announcement date from Compustat and analyst data from I/B/E/S. We obtain market data from CRSP and ownership data from Thomson 13F filings. We exclude articles with missing control variables in the respective quarter. Overall,

[20] While the aggregation on the sentence level allows us to split the sentiment based on *Add_Cont* and *Omit_Cont* topics, we lose some variation of the sentiment data compared to the word count proxy, as a sentence is either positive or negative regardless of how strong the sentiment in the sentence is, e.g., how many positive or negative words it contains.

[21] For any news story that mentions an entity, RavenPack provides a relevance score between 0 and 100 indicating how strongly related the entity is to the underlying news story, with higher values indicating greater relevance.

our sample consists of 6,540 earnings announcement and article pairs. Panel A of Table 3.1 outlines the sample selection process; Panel B and C show the sample distribution across years and industries, respectively.

Table 3.1: Sample description

Panel A: Sample selection	N	
S&P500 firms between 2010 and 2016	661	
-Missing earnings announcement dates	(45)	
Sample firms	616	
-Firms without financial press coverage	(105)	
-Missing data	(13)	
Firms in sample	498	
Articles about sample firms	7,333	
-Missing SEC Press Releases	(75)	
-Aggregation of multiple articles	(454)	
-Missing data	(264)	
Final Sample: EA & Article pairs	**6,540**	
Panel B: Sample distribution per year		
Year	N	in percent (%)
2010	745	11.39
2011	721	11.02
2012	664	10.15
2013	671	10.26
2014	1,276	19.51
2015	1,282	19.60
2016	1,181	18.06
Total	*6,540*	*100*
Panel C: Sample distribution per industry		
Fama-French industry code (12 industries)	N	in percent (%)
(1) Consumer Non-Durables	708	10.83
(2) Consumer Durables	134	2.05
(3) Manufacturing	467	7.14
(4) Oil, Gas, and Coal Extraction (Energy)	350	5.35
(5) Chemicals and Allied Products	249	3.81
(6) Business Equipment	930	14.22
(7) Telephone and Television Transmission	337	5.15
(8) Utilities	155	2.37
(9) Wholesale, Retail, and Some Services	1,067	16.31
(10) Healthcare, Medical Equipment, and Drugs	481	7.35
(11) Finance	1,015	15.52
(12) Other (e.g., Hotels, Entertainment)	647	9.89
Total	*6,540*	*100*

Notes: This table provides details on the sample selection process (Panel A), sample distribution per year (Panel B), and sample distribution per industry (Panel C).

3.3.2 Topic coverage in earnings announcements and press articles

Based on the perplexity score, we estimate the LDA model (see Section 3.2.2) using our sample of 6,540 observations (earnings announcements and corresponding press articles) based on a total of 35 unique topics.[22] We label each topic based on its high-frequency words. To further validate the topic labels, we read articles with a high proportion of a topic and evaluate if the content fits our label. Additionally, we read random articles and check whether we can identify all topics, which are included in the article according to the topic model. Table 3.2 reports for each topic the topic label and the 15 most frequent words for each topic, respectively. While our topics are distinct with respect to the high-frequency words and their respective probabilities, some topics are similar in their interpretation. For example, we have three different topics that deal with the end of a fiscal quarter. They contain words like "end", "second (quarter)", or "fourth (quarter)". Because all three sets of words refer to the same business matter (end of a quarter), we combine these three topics into one topic group to facilitate our analyses. Overall, we manually aggregate our 35 topics into 18 groups. Table 3.2 shows the assignment of each of the 35 topics into the 18 topic groups.

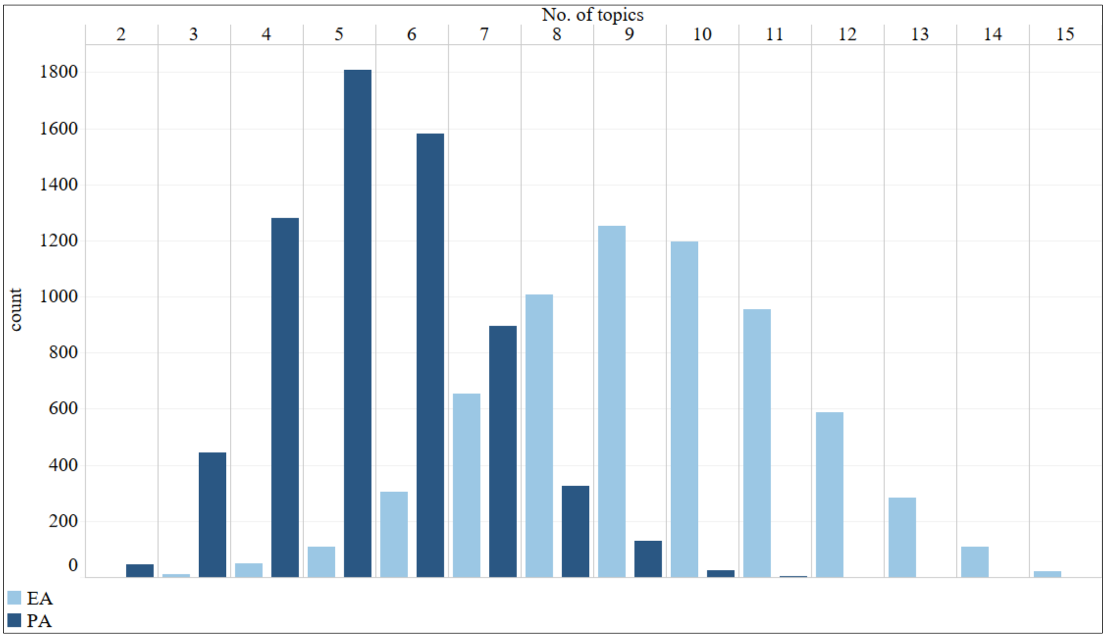


Figure 3.1: Topic distribution
This figure shows the number of topics in earnings announcements (EA) and corresponding press articles (PA) based on the LDA model.

---

[22] We manually compare the LDA model based on 30, 35, and 40 topics. While 30 topics do not seem to identify all important topics, the results of the output based on 40 topics are somewhat overlapping and reduces interpretability. Therefore, we conclude that 35 topics seems most suitable for our analysis.

Table 3.2: Topic groups

Topic No.	Topic	Group No.	Group	Words
1	Business Outlook	1	Business Outlook	expect, approximately, rate, include, result, base, guidance, estimate, cost, range, project, plan, end, follow, cash
2	Business Outlook			cost, high, low, price, volume, expect, increase, compare, result, demand, primarily, market, operation, continue, improve
3	CashFlow	2	CashFlow	flow, free_cash, operate, measure, organic, cash, cash_flow, growth, operating, eps, margin, business, income, statement, core
4	Compare	3	Compare	percent, increase, total, report, ago, expense, rate, expect, earning, statement, compare, approximately, growth, drive, basis
5	Earnings measures	4	Earnings	gaap, non, financial, measure, result, information, exclude, performance, management, provide, operating, income, net, expense, reconciliation
6	Earnings			net, income, expense, cash, total, tax, operating, loss, asset, other, cost, operate, end, interest, consolidated
7	EPS			cent, revenue, rise, earning, profit, earlier, expect, say, report, sale, fall, exclude, thomson_reuters, forecast, growth
8	Diluted Earnings			fiscal, earning, diluted, prior, operation, continue, end, expect, result, current, product, range, june, report, sale
9	Earnings Release	5	Earnings Release	financial, page, measure, gaap, period, consolidated, result, corporation, management, information, performance, non, statement, present, believe
10	Earnings Release			statement, forward_look, result, information, diluted, income, include, report, earning, net, financial, available, conference, factor, unaudited
11	Foreign Exchange	6	Foreign Exchange	rate, growth, u.s., foreign_currency, impact, exchange, basis, constant_currency, currency, report, prior, global, dollar, change, business
12	Growth	7	Growth	growth, market, business, product, continue, strong, new, customer, technology, sale, grow, demand, improve, global, increase
13	Banking Industry	8	Industry Specific	loan, net, increase, income, interest, expense, average, low, loss, compare, decrease, high, prior, total, capital
14	Insurance			loss, income, net, insurance, investment, business, operating, gain, financial, prior, result, premium, ratio, operate, tax
15	Utilities			earning, energy, high, include, increase, customer, operation, compare, cost, natural_gas, operate, rate, business, gas, project
16	Gas/Oil Industry			production, average, oil, natural_gas, drilling, net, price, operation, gas, activity, include, operate, oil_gas, project, cost
17	Retail Industry/ Brand			volume, increase, market, brand, high, growth, drive, low, decrease, impact, net, reflect, exclude, price total

18	Retail Industry			sale, store, increase, comparable, retail, home, end, new, brand, profit, earning, gross, open, period, inventory
19	Investment Activity			investment, asset, management, fund, client, income, revenue, net, performance, fee, equity, expense, market, include, firm
20	Contractors/ Service Providers			contract, program, increase, operate, compare, decrease, low, operating, primarily, segment, margin, high, system, backlog, volume
21	M&A	9	M&A	acquisition, cost, relate, business, integration, include, acquire, expense, merger, earning, benefit, exclude, transaction, health, performance
22	Adjusted earnings	10	Non GAAP	adjust, adjusted, gaap, eps, diluted, income, earning, report, measure, reconciliation, prior, continue, financial, result, operation
23	(adjusted) ebitda			ebitda, net, adjusted_ebitda, measure, cash, adjust, gaap, result, performance, loss, financial, interest, expense, debt, include
24	Third Quarter	11	Quarter End	month, end, september, period, compare, third, increase, net, october, income, primarily, respectively, relate, november, decrease
25	Forth Quarter			fourth, compare, december, increase, full, income, net, end, diluted, january, decrease, result, approximately, february, record
26	Second Quarter			second, june, compare, july, half, increase, month, end, income, result, period, august, decrease, approximately, diluted
27	RealEstate Investment	12	RealEstate Investment	property, real_estate, include, total, operating, lease, interest, debt, investment, market, portfolio, development, management, fund, gain
28	Risk	13	Risk	include, change, financial, result, business, statement, ability, impact, product, customer, market, risk, new, service, forward_look
29	Sales	14	Sales	sale, net, product, increase, u.s., earning, include, compare, decrease, primarily, rate, sales, change, high, worldwide
30	Revenue			revenue, service, increase, business, compare, period, total, services, customer, margin, growth, prior, software, operating, technology
31	Revenue			revenue, increase, high, operating, income, advertising, growth, include, reflect, digital, result, operate, network, international, drive
32	Segments	15	Segments	segment, profit, operating, business, financial, result, impact, cost, ago, earning, sale, change, include, item, period
33	Special Items	16	Special Items	special_item, fuel, group, expense, cost, exclude, revenue, net, include, unit, price, operating, total, hedge, impact
34	Statement	17	Statement	say, business, profit, u.s., not, new, fall, analyst, market, earlier, report, plan, expect, rise, inc.
35	Tax	18	Tax	tax, charge,relate, include, cost, result, prior, income, expense, impact, business, item, loss, exclude, certain

Notes: This table provides information about the topic aggregation into groups.

Figure 3.1 shows the distribution of topic amounts for both earnings announcements and corresponding press articles. It can be gleaned from these distributions that earnings announcements, on average, cover a larger number of topics, with a mean (median) number of 9.4 (9) topics per article and a fairly symmetric distribution. In comparison, press articles cover a much lower number of topics, with a mean (median) of 5.4 (5) topics per article. This finding is consistent with our maintained assumption that owing to space constraints and to an aggregation role, press articles are more focused on average. This assumption is also underscored by the difference in volumes: the average (median) word count of our sample earnings announcements is 2,852 (2,404) compared to 516 (455) for press articles, corresponding to roughly five times the volume.

Table 3.3, Panel A, reports details on the coverage of our 18 topic groups in earnings announcements and corresponding press articles and sheds light on the relative frequency or importance of these topics. The second column of Panel A shows that most prevalent topics are included in almost nine out of ten earnings announcements, with the top spot taken by "EarningsRelease" (89.4%), followed by "Earnings" (87.2%), and "IndustrySpecific" (86.3%). In comparison, topics such as "Statement" (15.3%), "SpecialItems" (16.4%), and "RealEstateInvestments" (17.0%) are least frequently included in earnings announcements and are to be found in less than one out of five firm announcements. All in all, these findings are in line, in particular, with the prominent role of earnings in corporate communication.

Column (3) of Panel A of Table 3.3 documents topic coverage in press articles. Consistent with a more focused, brief narrative, findings illustrate how journalists are more selective about the topics they choose to report on, compared to firms' earnings announcements. For example, only two topic categories are covered more frequently by journalists than by firms. This is the case for the category "Earnings", which receives even more frequent attention by the press than the substantial attention there already is in earnings announcements (94.8% versus 87.2%). Second, and not surprisingly, the topic group "Statement" appears in press articles in almost 95% of cases, compared to only 15.3% in earnings announcements. This finding reflects one archetypal source of information production by the press, which is supplementing their coverage of firm announcement narratives with statements from additional sources such as financial analysts, of senior firm managers (Call et al. (2018)).

Table 3.3: Topic Coverage in Earnings Announcements (EA) and Press Articles (PA)

Panel A: Topic coverage in EA and PA

	(1)		(2)		(3)		(4)	
Topic	Total sample (EA-PA pairs)		EAs including topic		PA including topic		Topic not featured in PA/ EA	
	N	Prop.	N	Prop.	N	Prop.	N	Prop.
BusinessOutlook	6,540	100.0%	4,289	65.6%	2,224	34.0%	1,758	26.9%
CashFlow	6,540	100.0%	2,403	36.7%	355	5.4%	4,052	62.0%
Compare	6,540	100.0%	2,758	42.2%	22	0.3%	3,767	57.6%
Earnings	6,540	100.0%	5,706	87.2%	6,200	94.8%	39	0.6%
EarningsRelease	6,540	100.0%	5,846	89.4%	175	2.7%	674	10.3%
ForeignExchange	6,540	100.0%	2,928	44.8%	1,380	21.1%	3,378	51.7%
Growth	6,540	100.0%	2,871	43.9%	2,030	31.0%	2,962	45.3%
IndustrySpecific	6,540	100.0%	5,641	86.3%	4,967	75.9%	492	7.5%
M&A	6,540	100.0%	2,169	33.2%	1,256	19.2%	3,984	60.9%
NonGAAP	6,540	100.0%	3,363	51.4%	492	7.5%	3,030	46.3%
QuarterEnd	6,540	100.0%	5,157	78.9%	2,765	42.3%	1,054	16.1%
RealEstateInvestments	6,540	100.0%	1,110	17.0%	412	6.3%	5,226	79.9%
Risk	6,540	100.0%	3,916	59.9%	490	7.5%	2,477	37.9%
Sales	6,540	100.0%	5,358	81.9%	3,913	59.8%	845	12.9%
Segments	6,540	100.0%	2,629	40.2%	962	14.7%	3,623	55.4%
SpecialItems	6,540	100.0%	1,074	16.4%	676	10.3%	5,118	78.3%
Statement	6,540	100.0%	998	15.3%	6,181	94.5%	330	5.0%
Tax	6,540	100.0%	3,351	51.2%	914	14.0%	2,971	45.4%

Panel B: Topic coverage in PA

Topic	PAs including topic						PAs without topic	
	(1)		(2)		(3)		(4)	
	Total		Topic in EA (*Emphasis*)		Topic not in EA (*Add_Cont*)		Topic in EA (*Omit_Cont*)	
	N	Prop.	N	Prop.	N	Prop.	N	Prop.
BusinessOutlook	2,224	100.0%	1,731	77.8%	493	22.2%	2,558	59.6%
CashFlow	355	100.0%	270	76.1%	85	23.9%	2,133	88.8%
Compare	22	100.0%	7	31.8%	15	68.2%	2,751	99.7%
Earnings	6,200	100.0%	5,405	87.2%	795	12.8%	301	5.3%
EarningsRelease	175	100.0%	155	88.6%	20	11.4%	5,691	97.3%
ForeignExchange	1,380	100.0%	1,146	83.0%	234	17.0%	1,782	60.9%
Growth	2,030	100.0%	1,323	65.2%	707	34.8%	1,548	53.9%
IndustrySpecific	4,967	100.0%	4,560	91.8%	407	8.2%	1,081	19.2%
M&A	1,256	100.0%	869	69.2%	387	30.8%	1,300	59.9%
NonGAAP	492	100.0%	345	70.1%	147	29.9%	3,018	89.7%
QuarterEnd	2,765	100.0%	2,436	88.1%	329	11.9%	2,721	52.8%
RealEstateInvestments	412	100.0%	208	50.5%	204	49.5%	902	81.3%
Risk	490	100.0%	343	70.0%	147	30.0%	3,573	91.2%
Sales	3,913	100.0%	3,576	91.4%	337	8.6%	1,782	33.3%
Segments	962	100.0%	674	70.1%	288	29.9%	1,955	74.4%
SpecialItems	676	100.0%	328	48.5%	348	51.5%	746	69.5%
Statement	6,181	100.0%	969	15.7%	5,212	84.3%	29	2.9%
Tax	914	100.0%	696	76.1%	218	23.9%	2,655	79.2%

Notes: This table shows the topic comparison for press articles (PA) and earnings announcements (EA) in Panel A and topic coverage in press articles in Panel B. The proportion of *Omit_Cont* in Panel B is calculated using the N of *Omit_Cont* (column (4)) scaled by the N of EAs including topic in Panel A (column (2)) (e.g., 2.558/4.289 for BusinessOutlook).

The "Earnings" and "Statement" topics rank first and second, respectively, within the press articles category, followed by industry-specific information ("IndustrySpecific", 75.9%). On the lower end of the distribution, we find topics that play a fairly minor role for press coverage of earnings announcements: "Compare" (0.3%), "EarningsRelease" (2.7%), and "CashFlow" (5.4%).

3.3.3 Information production in press articles

3.3.3.1 Adding and omitting topics

Panel B of Table 3.3 reports descriptive evidence for two of our four information production measures, the proportion of topics added in press articles (*Add_Cont*, Column (3)), and of topics omitted (*Omit_Cont*, Column (4)), respectively. Also, it reports the number and proportion of topics that are jointly covered in both documents (Column (2)). These statistics shed additional light on the topic coverage by journalists, as documented in Panel A. We illustrate this in the following by way of example for the first topic, "BusinessOutlook". Out of all 6,540 earnings announcements, firms include the topic "BusinessOutlook" in 65.6% of the cases (Column (2), Panel A of Table 3.3), rendering this topic a fairly frequent content of firms' earnings-related disclosures. Journalists, in turn, choose to cover the "BusinessOutlook" topic in 2,224 of the corresponding press articles (Column (3), Panel A of Table 3.3). Journalists' topic coverage is further disentangled in Panel B of Table 3.3, which reveals that out of the 2,224 press articles covering "BusinessOutlook", 77.8% represented "joint coverage", i.e., instances where "BusinessOutlook" was also a theme in the earnings announcements. Therefore, in the remaining 22.2% (493) articles, journalists choose to include information on the business outlook although no such topic information was reported in the corresponding earnings announcement, consistent with producing editorial content by adding new topics to their coverage (*Add_Cont*). In comparison, journalists choose to omit the topic in their articles much more frequently than adding it. According to Column (4) in Table 3.3, Panel B, journalists choose not to cover the topic „BusinessOutlook" in 59.6% of the cases, when it was discussed in the earnings announcement.

Turning to our first measure of information production, Figure 3.2 shows the distribution of *Add_Cont* topics for all press articles. The figure shows that information production via adding topics happens very frequently, as more than nine out of ten press articles (91.2%) add this type of information. In most cases, press articles add one or two topics.

On average, the press articles provide 1.5 additional topics compared to the earnings announcement. Topics frequently added are “Earnings” (795 articles), “Growth” (707 articles), “BusinessOutlook” (493 articles), and “IndustrySpecific” (407 articles) as noted in Panel B of Table 3.3. The most frequent “additional topic” is the “Statement” category, as a total of 5,212 press articles (84.3%) augment their narrative by adding this topic. It reflects journalists (privately) interacting with firm insiders or knowledgeable experts such as analysts to provide additional background information about the earnings announcement (Li (2014)). The prominence of the “Statement” topic complements the survey paper of Call et al. (2018), in which journalists state firm executives and analysts as important sources for their work. Table 3.A2 (Appendix) provides examples of such statements in financial press articles.

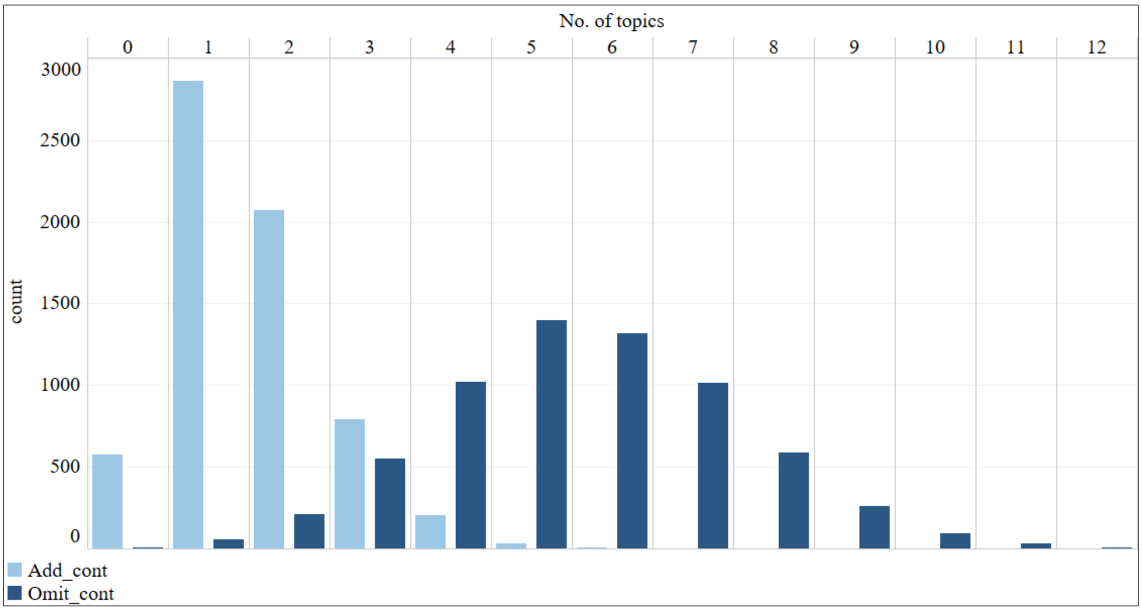


Figure 3.2: Distribution of added and omitted topics in press articles
This figure shows the number added (*Add_Cont*) and omitted topics (*Omit_Cont*) in press articles (PA) based on the LDA model.

With respect to our second measure of information production, *Omit_Cont*, Figure 3.2 documents that press articles on average leave out 5.5 topics that have been included in the corresponding earnings announcements. Panel B of Table 3.2 sheds light on which topics are thus considered most “irrelevant” in our sample. Accordingly, the topic “EarningsRelease” stands out, as it is covered in 5,846 earnings announcements, while rarely in the corresponding articles, which chose to exclude this topic in 5,691 cases. This is not surprising, as this topic features mostly technical information from the earnings announcement, such as the page number and the table layout of the balance sheet and income statement. Other frequently omitted topics, ranked from No. 2 to No. 5, are “Risk” (covered in 3,916 earnings announcements, omitted in 3,573 press articles), “NonGAAP” (covered in 3,363 earnings announcements, omitted in 3,018 press articles), and “Tax” (covered in 3,351 earnings announcements, omitted in 2,655 press articles).

3.3.3.2 Emphasis

Our third measure of information production, *Emphasis*, captures the relative significance that journalists assign to a specific topic, thereby choosing to (de-)emphasize that topic and thus conveying to readers a signal about relative importance, e.g., by adding additional analysis on that topic or emphasizing the relevance of a topic. The benchmark for relative significance is the corresponding earnings announcement. *Emphasis* for one particular topic is therefore defined as the topic proportion in the press article less the topic proportion in the earnings announcement.

By definition, *Emphasis* can only be computed for topics covered in both the earnings announcement and the press article. Therefore, our analysis of this third production measure anchors on the subsample of “jointly covered topics”, as reported in Column (2), Panel B of Table 3.3, which we further reduce to 15 topics.[23] Figure 3.3 plots the distribution of *Emphasis* for 15 topics using all respective joint pairs of earnings announcements and press articles.[24] We present the distribution per topic for two reasons: first, it highlights the on average higher emphasis of articles on the topics earnings, business outlook, and sales as well as the de-emphasis of the risk and non-GAAP topics. The t-statistics provided in Table 3.A3 (Appendix) support this notion of (de-)emphasis. Interestingly, *Emphasis* is not significantly different from zero for “Tax” and “RealEstateInvestment”, which suggests that journalists do not produce additional information through the emphasis channel for these topics. Second, it highlights the cross-sectional variation in *Emphasis*. For example, while the “Sales” topic is emphasized more on average, this positive mean is driven by a rather long tail of a few articles significantly focusing on sales. The significant cross-sectional variation is important since it hints at the possibility that information production is highly contextual. The “IndustrySpecific” topic exemplifies this contextuality. The mean emphasis of this topic is close to zero, yet it has a wide

[23] We exclude “Statement”, “EarningsRelease”, and “QuarterEnd” from the *EMPHASIS* analysis. Statement is mainly an added topic, while articles nearly completely omit EarningsRelease. Therefore, both topics are not suited for this analysis. Additionally, we exclude “QuarterEnd”, because this topic largely features technical information about the reporting period.

[24] Corresponding descriptive data is reported in the Appendix in Table A3.

distribution of cases where the industry topic is significantly de-emphasized, and a similar number of cases significantly increase the emphasis on this topic.

For the sake of brevity, we confine our discussion to topics with distinct distributions of *Emphasis*, in terms of mean (different from zero) and symmetry of the distribution. On this note, coverage of firm performance deserves a closer look. For one thing, the topic "Earnings" receives far more relative attention / emphasis in press articles, with a mean of 0.24, and shows a long leftward tail indicating that few press articles also quite drastically de-emphasize the discussion of firm earnings. First and foremost, this finding resonates with the pivotal role of earnings in firm valuation and with journalists catering to demands of their readership to focus on firm performance. In contrast, however, the opposite appears to hold for "NonGAAP" earnings, which receive relatively less attention by the press, with a mean of -0.097. This finding is somewhat surprising given the important role of pro forma earnings, and, more notably, evidence of potentially opportunistic use of these metrics to overstate firm performance (see, e.g. Bhattacharya et al. (2007)). There are at least two, non-exclusive explanations. First, firms that choose to report non-GAAP earnings are required by Regulation G to provide additional information, e.g., on the predictive ability of these metrics and reconciliations to GAAP earnings. Journalists potentially aggregate this information. Second, however, journalists may as well coin non-GAAP earnings as GAAP earnings, which refers to the "Earnings" topic. In that case, journalists potentially leave out information important to investors / their audience, as non-GAAP earnings potentially exhibit less reliability and more bias.

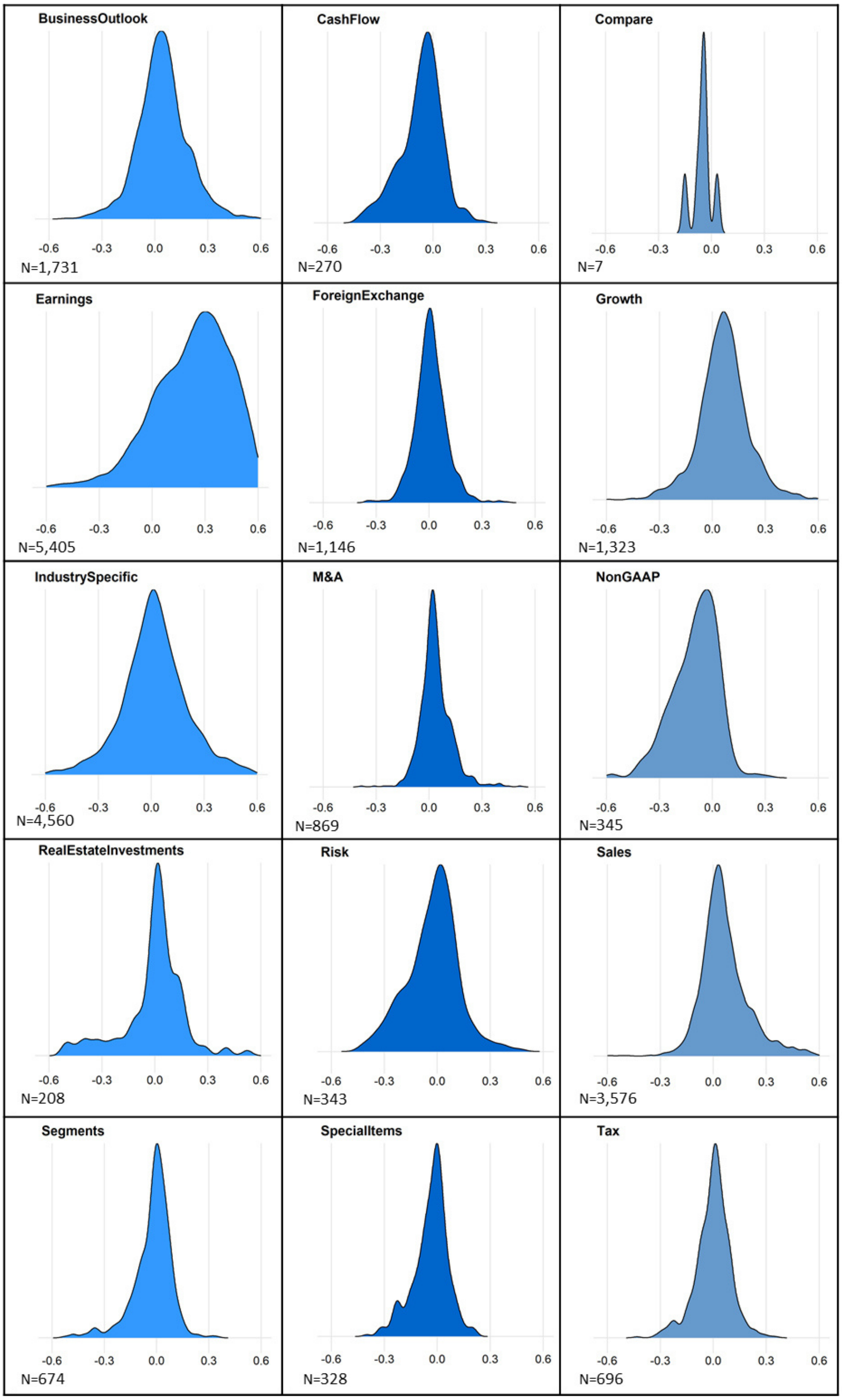


Figure 3.3: Density of Emphasis distribution
This figure shows the density plots for the *Emphasis* of different topics. It only includes article/earnings announcement pairs, which both feature the topic (Table 3.A3 Appendix). A positive value means the press emphasizes the topic, while a negative sign shows it understates the topic (N 6,540).

3.3.3.3 Sentiment

Our fourth and final measure of information production, *Sentiment*, captures not *what* topic journalists choose (not) to cover, but *how* they do so, adding a linguistic angle to our analyses. In a first step, we examine the sentiment of the earnings announcements and corresponding press articles of the level of the whole document. We compute the sentiment on the document level based on the positive and negative word counts. Figure 3.4, Panel A shows the boxplots of the sentiment distribution. Generally, the articles have a more negative sentiment (mean:-0.17, median:-0.20) than the corresponding earnings announcements (mean:-0.02, median:-0.04). Both the mean and the median of press articles are significantly more negative than the earnings announcements on the 1% level.

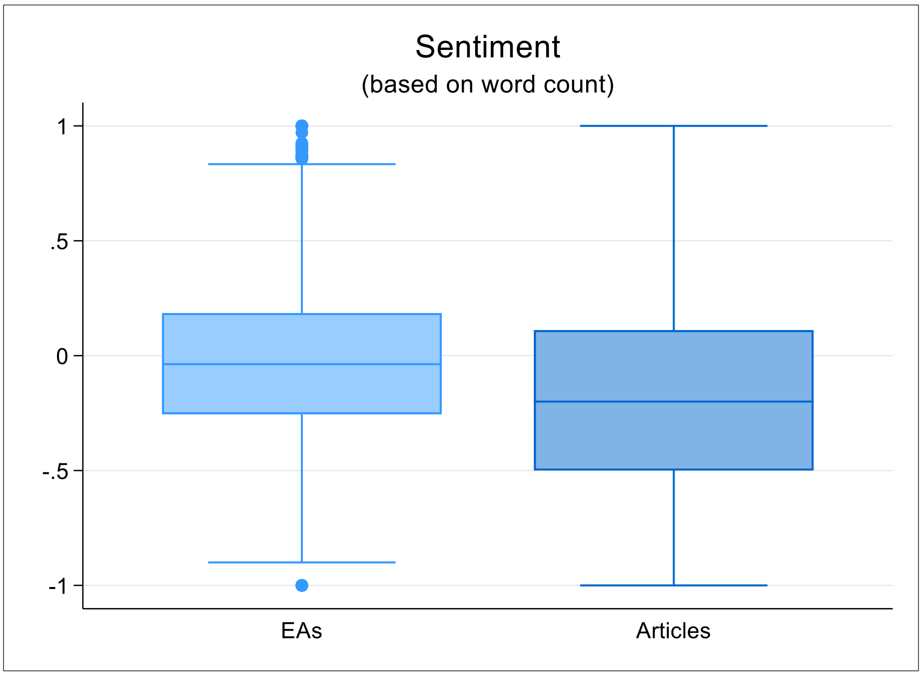
As topic coverage differs between press articles and earnings announcements, we further distinguish *Sentiment* for joint topics, i.e., for topics in the press article which are also featured in the corresponding earnings announcements, and for topics added by journalists (see Section 3.3.3.1) in a second step. To make this distinction, we need to aggregate the sentiment on the sentence level. Figure 4, Panel B shows the sentiment on the sentence level and the split into sentiment based on added and joint topics. Comparable to the word count level, the overall article sentiment is negative and significantly different from zero at the 1% level (mean:-0.09). The mean sentiment of jointly covered topics (*Sent_Joint_Topics*) shows a negative mean of -0.10, which is significantly different from zero at the 1% level. Additionally, the mean sentiment of topics that journalists choose to add relative to the firm's earnings announcements (*Sent_Add_Topics*) is -0.04 and also significantly different from zero at the 1% level. The difference between the mean and median (0 vs. -0.09) of *Sent_Add_Topics* and *Sent_Joint_Topics* is significant at the 1% level. Thus, coverage of firms' earnings announcements by journalists, on average, is more negative, or critical, in tone than coverage of additional topics.

On the topic level, the median sentiment of most topics is close to zero.[25] For example, "Statement", which is mainly an added topic, has a mean and median close to zero and a rather symmetric distribution around its median. This distribution suggests that journalists equally feature positive and negative firm executive or expert quotes. The overall

[25] Figure 3.A4 (Appendix) shows the distribution of the sentiment for the various topics.

sentiment of "BusinessOutlook" has a mean of -0.13 (median: 0). The sentiment of this topic is more negative when it is a joint topic instead of an added topic (-0.14 vs. -0.08).

Panel A: Sentiment based on the word count



Panel B: Sentiment based on the sentence level

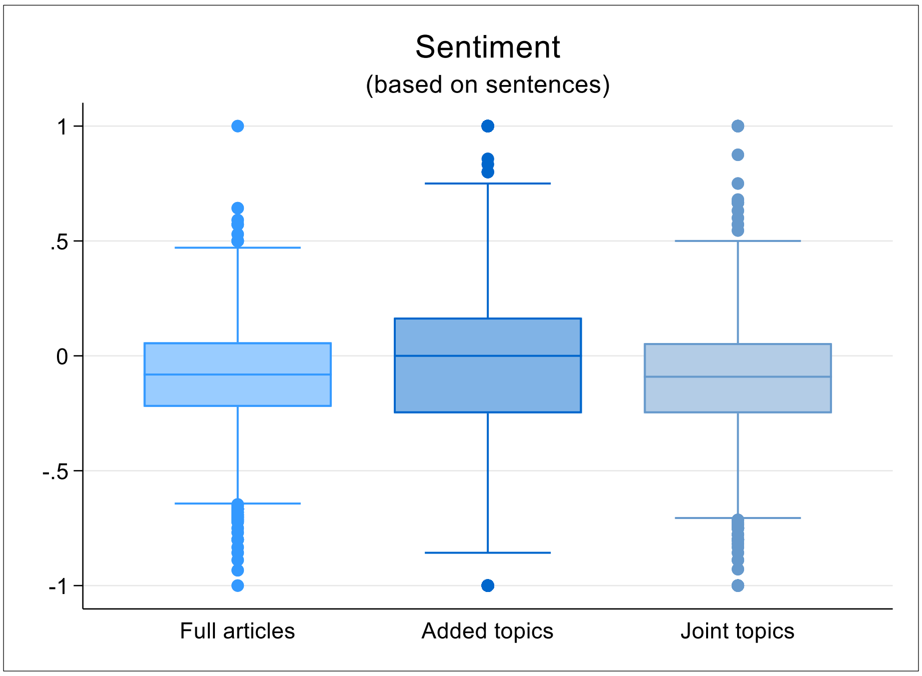


Figure 3.4: Sentiment distribution

Panel A shows the box plots for the sentiment of the earnings announcements and press articles. The difference between the median sentiment of earnings announcements (median=-0.037) and media articles (median=-0.200) is significantly different at the 1% level based on a Wilcoxon signed-ranks test (z-value=24.856). Panel B shows the box plots for the sentiment of the topics on the sentence level. The difference between the median sentiment of added (median=0) and joint topics (median=-0.090) is significantly different at the 1% level based on a Wilcoxon signed-ranks test (z-value=11.286).

3.4 How do journalists help investors analyze firms' earnings news?

Our second set of empirical tests answers our second research question: how do journalists help investors analyze firms' earnings news? As noted in the introduction, Guest (2018) illustrates that information production by the financial press creates value to investors, as market reactions to earnings announcements are larger when these earnings announcements receive concurrent press article coverage, and these market reactions increase with the proportion of editorial content contained in the press article. However, Guest (2018) illustrates a measure of editorial content, dissimilarity between earnings announcements and articles, which is fairly opaque with respect to how journalists create value or content. This is where our analyses kick in: Using our different measures of journalists' information production that we glean from topic analyses (see Section 3), we investigate how market reactions to earnings announcements vary with the amount of these different sources of information production. In the remainder of this section, we outline our research design (Section 3.4.1) and present our empirical findings from market reaction tests (Section 3.4.2).

3.4.1 Research Design

3.4.1.1 Model

To test our expectations regarding the information production role of the media, we estimate the following model:

$$(1) \quad MABHR_{it} = \beta_0 + \beta_1 Surprise + \beta_2 Production + \beta_3 Surprise\ x\ Production + \beta_4 Controls + \beta_5 Firm\ FE + \beta_6 Year\_Quarter\ FE + \varepsilon_{it}$$

We estimate Eq. (1) using OLS regression, quarter-year and firm fixed effects, and standard errors clustered at the firm level. Our dependent variable *MABHR* is the three-day market-adjusted buy-and-hold return centered on the earnings announcement day. We use the NYSE, AMEX, Nasdaq value-weighted market index for the market adjustment (Guest (2018)). We measure earnings surprise (*Surprise*) as the difference between actual earnings per share and the last median analyst forecast before the earnings announcement day (both based on I/B/E/S), scaled by stock price on the forecast date (Huang et al. (2017)). The coefficient of *Surprise*, β_1, is the earnings response coefficient (ERC) and captures the markets' response to the unexpected portion of the reported earnings.

Production represents our various measures of information production by the financial press, which we introduced in Section 2.1., and for which we presented descriptive findings in Section 3. Specifically, we estimate different versions of our main model (Eq. (1)) for three of our information production measures: additional content (*Add_Cont*), omitted content (*Omit_Cont*), and sentiment (*Sentiment*). Continuous production and sentiment variables are standardized to facilitate the interpretation of coefficients.

We do not include our fourth information production measure, *Emphasis,* in the multivariate regressions for two reasons. First, *Emphasis* is measured on the topic level and unique for each topic, which does not allow a meaningful aggregation of the measure on the article level. Second, and more importantly, we are interested in how the market reaction varies based on different levels of information production. *Emphasis,* however, relies on joint topics and, therefore, on information, which journalists (partly) disseminate from the earnings announcements. Thus, this proxy does not allow the clear identification of information production needed for the multivariate analysis.

While *Production* captures the main effect of the three information production proxies, our main variable of interest in Eq. (1) is the interaction term *Surprise x Production*. The coefficient β_3 captures the conditional relationship between unexpected earnings and information production in financial press articles. This coefficient assumes values significantly different from zero if information production by the press helps investors analyze earnings new by firms (Guest (2018)).

3.4.1.2 Control variables

The vector *Controls* in Eq. (1) includes three sets of control variables, which respectively control for the information environment of the firm, for firm characteristics, and for information characteristics of the earnings announcement. To control for the information environment of the firm, we include into our vector of control variables three measures that proxy for firm coverage by the media and by analysts. *Ln_Analyst* is the natural logarithm of one plus the number of analysts that issue earnings forecasts. *Ln_Prior* is the natural logarithm of one plus the number of days with DowJones newswire coverage in the year before the earnings announcement. *Ln_EA_Press* is the natural logarithm of one plus the number of DowJones newswire articles about the firm on the day of the earnings announcement (Twedt (2016)).

Table 3.4: Summary statistics

	mean	p50	sd	min	p25	p75	max	count
MABHR	0.001	0.001	0.062	-0.472	-0.029	0.032	0.500	6,540
Surprise	0.001	0.000	0.004	-0.017	0.000	0.001	0.021	6,540
Added_Topics	0.715	0.679	0.990	-0.714	-0.172	1.534	3.366	6,540
Omitted_Topics	0.007	-0.038	0.997	-2.652	-0.703	0.703	2.923	6,540
Sent_Art	-0.182	-0.248	1.145	-2.195	-0.978	0.509	2.673	6,540
Sent_EA	0.002	-0.041	1.001	-3.016	-0.714	0.646	3.163	6,540
Ln_EA_Press	1.604	1.609	0.593	0.693	1.099	1.946	3.258	6,540
Bad_News	0.220	0.000	0.414	0.000	0.000	0.000	1.000	6,540
Ln_Prior	4.254	4.205	0.572	2.996	3.861	4.605	5.645	6,540
Ln_Analyst	2.956	2.996	0.432	1.099	2.773	3.258	3.689	6,540
Ln_Mv	10.061	10.009	1.159	7.276	9.278	10.848	12.696	6,540
BTM	0.418	0.335	0.351	-0.418	0.190	0.570	1.694	6,540
Ret_Sd	0.077	0.067	0.039	0.028	0.050	0.091	0.249	6,540
Ln_Emp	3.611	3.593	1.212	1.104	2.698	4.486	6.075	6,540
Ln_Owner	2.868	2.805	1.720	0.056	1.478	3.929	7.279	6,540
Inst_Ownership	0.698	0.772	0.273	0.000	0.652	0.865	1.099	6,540
Ln_8K	2.418	2.565	0.846	0.000	2.197	2.890	3.807	6,540
Ln_EAD	6.169	6.284	0.763	4.159	5.687	6.807	7.294	6,540
Comp	-0.012	-0.028	1.156	-2.718	-0.780	0.730	3.143	6,540
Sent_Add_Topics	-0.002	0.082	0.996	-2.308	-0.516	0.480	2.471	5,068
Sent_Joint_Topics	-0.028	0.000	1.009	-3.928	-0.659	0.621	4.791	5,068

Notes: This table provides summary statistics for the variables. Continuous variables are winsorized at the 1% level. Variables comprise MABHR (market-adjusted buy-and-hold return), Surprise (earnings surprise), Added_Topics (proportion added topics), Omitted_Topics (proportion omitted topics), Sent_Art (sentiment press articles), Sent_EA (sentiment earnings announcements), Ln_EA_Press (number of press articles about the earnings announcement), Bad_News (negative earnings surprise indicator), Ln_Prior (number of days with prior press coverage), Ln_Analyst (analyst following), Ln_Mv (firm size), BTM (Book-to-market), Ret_Sd (return volatility), Ln_Emp (number of employees), Ln_Owner (number of owners), Inst_Ownership (institutional ownership), Ln_8K (number of prior 8Ks), Ln_EAD (competing earnings announcements), Comp (principal component analysis of textual features of the earnings announcement), Sent_Add_Topics (sentiment added topics), and Sent_Joint_Topics (sentiment joint topics). All variables are defined in Table 3.A1 (Appendix).

To control for general firm characteristics, we further include into *Controls* several variables, all of which we measure at the quarter end of the earnings announcement (Fang and Peress (2009); Bushee et al. (2010); Drake et al. (2014); Guay et al. (2016); Guest (2018)). We use the natural logarithm of the market value of equity (*Ln_Mv*) as a proxy for firm size. *BTM* is the book value of equity scaled by the market value of equity. *LN_8K* is the natural logarithm of the number of 8-Ks filed in the year before the earnings announcement. *Inst_Ownership* is the percentage of institutional ownership on the most recent date available in the three months before the earnings announcement. We measure

Ln_Emp as the natural logarithm of one plus the number of employees, whereas *Ln_Owner* is the natural logarithm of one plus the number of owners at the fiscal year-end. *Ret_Sd*, a proxy for uncertainty, is the standard deviation of monthly returns over 12 months prior to the earnings announcement.

As a last set of control variables, we include variables directly related to the current earnings announcement. *Ln_EAD* is the natural logarithm of one plus the number of firms announcing their earnings in the *MABHR* window. *Bad_News* is an indicator variable set to one if earnings surprise is negative. *Comp* is the result of a principal component analysis of the textual features of the earnings announcement press release (Guest (2018)). For this analysis, we include readability, hard information, and specific entities. We measure readability as the product of minus one and the Gunning Fog index (Li (2008)) of the earnings announcement. Hard information refers to the count of numbers in the press release text. Following Blankespoor (2019), we exclude years, dates, section numbers, and descriptions. We identify specific entities mentioned in the earnings announcement using the spaCy Named Entities Recognition (NER) (Hope et al. (2016)). We winsorize all control variables at the 1% and 99% level due to outliers. Table 3.4 reports summary statistics for all variables.

3.4.2 Empirical Findings

3.4.2.1 Added topics

For our market tests, we create *Added_Topics* as a measure of information production via adding new topics to press articles. *Added_Topics* adds up the proportion of each *Add_Cont* topic in a press article and thus captures how much of the article is based on information production. To further identify the subset of press articles that particularly focus on adding content, we create the binary variable *Added_Topics_Quart,* which assumes the value of one for articles belonging to the top 25% with respect to the proportion of added topics and zero otherwise.

Table 3.5 presents results from estimating our main model (Eq. (1)) for our measures of added content. As expected, the coefficient on *Surprise* is positive and significant in all model specifications (1% level), indicating information content of earnings surprises. The main effect of the proportion of added topics yields no significant results. The coefficient of our main variable of interest, the interaction term of *Surprise* and *Added_Topics*, assumes a significant (10% level) positive value of 0.557 (Column (2)), suggesting that the more original content the article contains, the stronger is the investor reaction to the unexpected proportion of the reported earnings. This finding is consistent with journalists helping market participants to understand and interpret the information contained in earnings announcements.

Table 3.5: Market responses and information production by journalists: Added topics

	Added Topics		Top quartile (Added Topics)	
	(1)	(2)	(3)	(4)
	MABHR	MABHR	MABHR	MABHR
Surprise	2.257***	1.832***	2.262***	1.852***
	(5.62)	(3.80)	(5.66)	(4.68)
Added_Topics	-0.001	-0.002		
	(-0.89)	(-1.24)		
Surprise x Added_Topics		0.557*		
		(1.71)		
Added_Topics_Quart			-0.003	-0.005*
			(-1.40)	(-1.91)
Surprise x Added_Topics_Quart				1.512*
R^2	0.189	0.190	0.189	0.191
N	6,540	6,540	6,540	6,540
Cluster	Firm	Firm	Firm	Firm
FEs	Firm, YQ	Firm, YQ	Firm, YQ	Firm, YQ
Controls	Yes	Yes	Yes	Yes

Notes: This table shows results from estimating OLS regressions using the market-adjusted buy-and-hold return (MABHR) as the dependent variable. Variables comprise Surprise (earnings surprise), Added_Topics (proportion added topics), and Added_Topics_Quart (indicator variable for top 25% of Added_Topics). Unreported control variables comprise Ln_EA_Press (number of press articles about the earnings announcement), Bad_News (negative earnings surprise indicator), Ln_Prior (number of days with prior press coverage), Ln_Analyst (analyst following), Ln_Mv (firm size), BTM (Book-to-market), Ret_Sd (return volatility), Ln_Emp (number of employees), Ln_Owner (number of owners), Inst_Ownership (institutional ownership), Ln_8K (number of prior 8Ks), Ln_EAD (competing earnings announcements), and Comp (principal component analysis of textual features of the earnings announcement). All variables are defined in Table 3.A1 (Appendix). ***, **, * indicate statistical significance at the 1%, 5%, and 10% level, respectively, using two-tailed tests and standard errors clustered as indicated.

From Column (4), it can be gleaned that this positive association between added topics in press articles and the ERC increases with the magnitude of information production, as the coefficient on the interaction term *Surprise x Added_Topics_Quart* is larger in magnitude compared to Column (2), with a value of 1.512. This finding demonstrates that articles in the highest *Added_Topics* quartile lead to an ERC that is 82% larger compared to earnings announcements with corresponding press articles that contain fewer additional content (i.e., 1.512/1.852).

3.4.2.2 Omitted topics

Similar to the *Added_Topics* variable (Section 3.4.2.1), we create *Omitted_Topics,* which measures information production via omitting content from the earnings announcement on the article level. For this proxy, we aggregate the earnings announcement topic proportions of all *Omit_Cont* topics. i.e., topics that firms feature in their disclosure, but which journalists decide not to report in the corresponding articles. As in the previous section, we add an indicator variable (*Omitted_Topics_Quart*), assuming one for the top quarter of *Omitted_Topics* articles and zero otherwise. Because omission of topics may emanate from article space constraints rather than from a journalist's attempt to leave out irrelevant content, we add the proportion of added topics (*Added_Topics*) as an additional control variable to our baseline regression model (Eq. (1)).

Table 3.6 reports the results from estimating Eq. (1) using *Omitted_Topics* as the production variable of interest. As in Table 3.5, *Surprise* is positive and highly significant through all model specifications. In contrast, the main effect for *Omitted_Topics* is insignificant and close to zero. Our main variable of interest is *Surprise x Omitted_Topics* in Column (2). This interaction term is significant at the 5% level and obtains a coefficient of 0.830. This result suggests that by deciding not to disseminate parts of the content from the earnings announcement, the financial press helps investors focus on value relevant information, as reflected in the magnified investors' response to earnings news.

Columns (3) and (4) introduce the indicator variable for press articles with a particularly high proportion of omitted content from the earnings announcements. The interaction term *Surprise x Omitted_Topics_Quart* is significant at the 5% level and has a magnitude of 2.128. Therefore, articles omitting most (highest quartile) content from the corresponding earnings announcement prompt an ERC that is about 110% larger compared to articles featuring more content from the earnings announcement (i.e., 2.128/1.938). Consequently, our results suggest that a more focused reporting, i.e., omitting more potentially irrelevant topics, which might distract investors, induces a more efficient response to earnings.

Table 3.6: Market responses and information production by journalists: Omitted Topics

	Omitted Topics		Top quartile (Omitted Topics)	
	(1) MABHR	(2) MABHR	(3) MABHR	(4) MABHR
Surprise	2.257***	2.459***	2.257***	1.938***
	(5.62)	(5.90)	(5.62)	(5.04)
Omitted_Topics	0.001	0.000		
	(0.68)	(0.04)		
Surprise x Omitted_Topics		0.830**		
		(2.45)		
Omitted_Topics_Quart			-0.001	-0.003
			(-0.39)	(-1.20)
Surprise x Omitted_Topics_Quart				2.128**
				(2.25)
Added_Topics	-0.001	-0.001	-0.001	-0.001
	(-0.96)	(-1.04)	(-0.86)	(-0.89)
R^2	0.189	0.191	0.189	0.191
N	6,540	6,540	6,540	6,540
Cluster	Firm	Firm	Firm	Firm
FEs	Firm, YQ	Firm, YQ	Firm, YQ	Firm, YQ
Controls	Yes	Yes	Yes	Yes

Notes: This table shows results from estimating OLS regressions using the market-adjusted buy-and-hold return (MABHR) as the dependent variable. Variables comprise Surprise (earnings surprise), Omitted_Topics (proportion omitted topics), and Omitted_Topics_Quart (indicator variable for top 25% of Omitted_Topics). Unreported control variables comprise Ln_EA_Press (number of press articles about the earnings announcement), Bad_News (negative earnings surprise indicator), Ln_Prior (number of days with prior press coverage), Ln_Analyst (analyst following), Ln_Mv (firm size), BTM (Book-to-market), Ret_Sd (return volatility), Ln_Emp (number of employees), Ln_Owner (number of owners), Inst_Ownership (institutional ownership), Ln_8K (number of prior 8Ks), Ln_EAD (competing earnings announcements), and Comp (principal component analysis of textual features of the earnings announcement). All variables are defined in Table 3.A1 (Appendix). ***, **, * indicate statistical significance at the 1%, 5%, and 10% level, respectively, using two-tailed tests and standard errors clustered as indicated.

3.4.2.3 Sentiment

As outlined in section 3.2.1, we measure the sentiment on the article level (*Sent_Art*) as the difference between positive and negative words, scaled by the sum of positive and negative words. The identification of positive and negative words is based on the Loughran and McDonald (2011) word lists. We split *Sent_Art* into two components: 1) the sentiment contained in those parts of the press article which cover topics in the earnings announcements, i.e., joint topics (*Sent_Joint_Topics),* and 2) the sentiment in the press article narratives that contain added topics (*Sent_Add_Topics).* We measure *Sent_Add_Topics* (*Sent_Joint_Topics)* on the sentence level, as the difference between positive and negative added topic (joint topic) sentences scaled by all added topic (joint topic) sentences in an article. We classify a sentence as positive (negative) if the number of positive words is larger (smaller) than the number of negative words.

For our analyses of *Sentiment* as the production variable of interest, we add to our baseline regression model (Eq. (1)) the sentiment of the earnings announcement (*Sent_EA*) as a control variable, as journalists might disseminate content from the earnings announcement and alter the sentiment of this information, but investors still might react to the sentiment of the original disclosure. We measure the tone of the earnings announcement as the difference between positive and negative words based on Loughran and McDonald (2011) scaled by the sum of positive and negative words.

Table 3.7 reports the results of estimating Eq. (1) using our sentiment variables. In all model specifications, *Sent_EA* obtains insignificant coefficients. Columns (1) and (2) examine the sentiment on the article level. The main effect of the sentiment is highly statistically significant (1% level) with a coefficient of 0.009 (Column (1)). Column (2) adds the interaction term of the article sentiment and the earnings surprise. *Surprise x Sent_Art* is significant at the 5% level with a coefficient of 0.515. This finding suggests that investors not only value *what* journalists choose to report on, but also *how* journalists do so.

In Column (3), *Sent_Art* is split into the sentiment of added topics (*Sent_Add_Topics)* and joint topics (*Sent_Joint_Topics)*. Both variables obtain significant coefficients at the 1% level. However, the coefficient for *Sent_Joint_Topics* is larger than the coefficient *Sent_Add_Topics*. An F-test shows that the difference between the two coefficients is statistically significant (p-value=0.004, F-statistic=8.38). Column (6) reports results for the full model, including the interaction terms of *Surprise* and the sentiment of added and joint topics. In contrast to the main effect, only the interaction term *Surprise x Sent_Joint_Topics* obtains a significant coefficient (10% level).

Table 3.7: Market responses and information production by journalists: Sentiment analysis

	Full article		Joint and Added topics			
	(1)	(2)	(3)	(4)	(5)	(6)
	MABHR	MABHR	MABHR	MABHR	MABHR	MABHR
Surprise	2.123***	2.499***	2.435***	2.489***	2.847***	2.856***
	(5.40)	(5.87)	(6.38)	(6.81)	(6.65)	(6.76)
Sent_Art	0.009***	0.008***				
	(10.12)	(9.47)				
Surprise x Sent_Art		0.515**				
		(1.99)				
Sent_Add_Topics			0.003***	0.003***	0.003***	0.003***
			(3.51)	(3.13)	(3.52)	(3.27)
Sent_Joint_Topics			0.008***	0.008***	0.007***	0.007***
			(6.97)	(6.97)	(6.26)	(6.27)
Surprise x Sent_Add_Topics				0.181		0.062
				(0.71)		(0.25)
Surprise x Sent_Joint_Topics					0.559*	0.547*
					(1.94)	(1.91)
Sent_EA	0.001	0.001	0.000	0.000	0.000	0.000
	(0.68)	(0.65)	(0.05)	(0.05)	(0.01)	(0.01)
R^2	0.207	0.208	0.228	0.228	0.229	0.229
N	6,540	6,540	5,068	5,068	5,068	5,068

H_0: Sent_Add_Topics = Sent_Joint_Topics; F-statistic=8.38, p-value=0.004

H_0: Sent_Add_Topics + Surprise x Sent_Add_Topics = Sent_Joint_Topics + Surprise x Sent_Joint_Topics; F-statistic=1.65, p-value= 0.199

Cluster	Firm	Firm	Firm	Firm	Firm	Firm
FEs	Firm, YQ	Firm, YQ	Firm, YQ	Firm, YQ	Firm, YQ	Firm, YQ
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table shows results from estimating OLS regressions using the market-adjusted buy-and-hold return (MABHR) as the dependent variable. Variables comprise Surprise (earnings surprise), Sent_Art (sentiment press articles), Sent_EA (sentiment earnings announcements), Sent_Add_Topics (sentiment added topics), and Sent_Joint_Topics (sentiment joint topics). Unreported control variables comprise Ln_EA_Press (number of press articles about the earnings announcement), Bad_News (negative earnings surprise indicator), Ln_Prior (number of days with prior press coverage), Ln_Analyst (analyst following), Ln_Mv (firm size), BTM (Book-to-market), Ret_Sd (return volatility), Ln_Emp (number of employees), Ln_Owner (number of owners), Inst_Ownership (institutional ownership), Ln_8K (number of prior 8Ks), Ln_EAD (competing earnings announcements), and Comp (principal component analysis of textual features of the earnings announcement). All variables are defined in Table 3.A1 (Appendix). ***, **, * indicate statistical significance at the 1%, 5%, and 10% level, respectively, using two-tailed tests and standard errors clustered as indicated.

Taken together, our results suggest that *how* journalists phrase their additional content matters: by "toning down" the topics circulated by firms in their earnings announcement, journalists appear to increase investors' ability to interpret the reported earnings. This is consistent with journalists going beyond the dissemination of these "joint" topics by providing useful analysis and discussion of firms' topics. In contrast, our findings do not suggest that sentiment produces value for original analysis by journalists, i.e., analyses that go beyond the firm's earnings announcements by adding topics. One explanation for this finding pertains to the nature of information journalists add to their articles. As outlined in section 3.3.3, the topic that WSJ journalists add most often is "Statement", i.e., direct or indirect quotes from firm executives or analysts. Adding this kind of information ties journalists to the wording of the quoted source and potentially limits their ability to alter the sentiment. In comparison, for joint topics, journalists have a broader range of how to phrase their narratives, which potentially facilitates deliberate information production via adding a specific tone.

3.5 Conclusion

In this paper, we investigate how journalists produce information in financial press articles and whether different forms of information production help investors analyze the earnings news. To investigate the information production, we examine the WSJ coverage of quarterly earnings announcements of S&P 500 firms between 2010 and 2016. Using a topic modeling approach (LDA), we model the content of financial press articles and the corresponding earnings announcements. The identification of topics enables us to develop four measures of information production: 1) adding content, 2) omitting content, 3), emphasizing content, and 4) applying sentiment to the content. In our first set of analyses, we provide a nuanced topic-based perspective on information production by the financial press, documenting that journalists make extensive, yet varying use of all four dimensions of information production. Our second set of analyses demonstrates that these information production activities appear to create value for investors. We find that both the activities of adding and omitting content in press articles are associated with larger investor responses to earnings news. For our information production measure "sentiment", we also find such a positive association, which is, however, confined to topics that are also covered in the earnings announcements. This result is consistent with journalists creating value for their audiences by "toning down" firms' narratives.

Our findings contribute to the very recent literature on the information production role of the financial press. We corroborate prior evidence that the press assumes such a role, and that it does so to the benefit of investors. Going beyond the prior literature, we are able to demonstrate *how* the press assumes its information role: by deliberately adding and omitting specific topics, and, with a smaller impact on investor reactions, by "toning down" firms' narratives. These insights increase our understanding of the economic role of the financial press as an intermediary of accounting information. At the same time, we provide a topic-based explanation for prior empirical findings that rest on potentially ambiguous measures of journalists' information production.

Our findings are subject to several limitations. First, the decision to cover a firm lies with the news outlet and is therefore endogenous. For example, journalists may be covering firms with specific topics, e.g., executive compensation, M&A activities, etc., or depending on the availability of new information, e.g., from interviews with executives. While this limitation is common in the financial press literature, this endogeneity concern limits our ability to draw causal inferences from our results. Second, our results rely on the construct validity of our production measures. These production measures derive from the LDA model, which is applied using various subjective inputs by the researchers. For example, we select the number of topics for which the algorithm estimates the LDA model. Also, the topic labels and, therefore, the topic groups are subject to our interpretation of the high-frequency words. Third, our sentiment analysis is based on the Loughran and McDonald (2011) dictionary. The authors developed this dictionary for 10-K disclosure narratives. While the dictionary should thus be suited for the analysis of earnings announcements, press articles represent a different kind of narrative, and a more general dictionary might be more suited for their analysis. Finally, our findings potentially do not generalize to less specialized news outlets and smaller firms with a different information environment than S&P 500 firms.

These limitations aside, our paper sheds a nuanced light on the information production role of the financial press and its potential usefulness to investors. It is up to future research to explore in more detail how this information production role varies across different news outlets, exploring further the press landscape. Also, given the structural changes in the production and dissemination of information brought about by digital transformation, it is an open question how the economic role of the financial press will evolve in the very near future.

3.6 Appendix

Appendix 3.A1: Variable definition

Panel A: Topics

Topic number	Topics	Topic number	Topics
1	BusinessOutlook	11	QuarterEnd
2	CashFlow	12	RealEstateInvestments
3	Compare	13	Risk
4	Earnings	14	Sales
5	EarningsRelease	15	Segments
6	ForeignExchange	16	SpecialItems
7	Growth	17	Statement
8	IndustrySpecific	18	Tax
9	M&A		
10	NonGAAP		

Panel B: Variables

Variable	Definition
MABHR	Market adjusted buy-and-hold return [-1/+1], adjusted for NYSE, AMEX, Nasdaq value-weighted market index
Surprise	Difference between actual earnings per share and the last median analyst forecast before the EAD (both based on IBES), scaled by stock price on the forecast date
Controls	
BTM	Book value of equity scaled by the market value of equity at the quarter-end [ceqq/(chsoq*prccq)]
Ln_Emp	Natural logarithm of 1 plus the number of employees at the fiscal-year end
Inst_Ownership	Percentage of institutional ownership on the most recent date available in three months prior to the earnings announcement. The variable is set to 0 if the firm is not featured in the database.
Ln_Mv	Natural logarithm of the market value of equity at the quarter-end [chsoq*prccq]
Ln_Owner	Natural logarithm of 1 plus the number of owners at the fiscal-year end
Ret_Sd	Standard deviation of monthly returns over 12 months prior to the earnings announcement
Ln_EA_Press	Natural logarithm of the number of DJ news articles about the firm on the earnings announcement day
Ln_Prior	Natural logarithm of the number of days with DJ press coverage in the year before the earnings announcement
Ln_Analyst	Natural logarithm the number of analysts with earnings forecasts
Ln_8K	Natural logarithm of the number of 8-Ks filed in the year before the earnings announcement
Bad_News	Indicator variable set to one if earnings surprise is negative

Ln_EAD	Natural logarithm of the number of Compustat firms announcing their earnings during the earnings announcement window (t-1/t+1)
Comp	First principal component of readability, hard information, and specific. Readability= -1*Gunning Fog index Hard information = (count of number in text/word count)*1000 Specific = (count of entities based on spacy classifier/word count)*1000
Production variables	
Add_Cont	Topic featured in the article but not in the earnings announcement
Omit_Cont	Topic featured in earnings announcement but not in the article
Emphasis	Difference between article topic proportion and earnings announcement topics proportion
Added_Topics	Proportion of all dropped topics in articles (standardized)
Added_Topics_Quart	Indicator variable set to one if article belongs to the top quartile of Added_Topics
Omitted_Topics	Proportion of all dropped topics in articles (standardized)
Omitted_Topics_Quart	Indicator variable set to one if article belongs to the top quartile of Omitted_topics
Tone variables	
Sent_Add_Topics	Difference between positive and negative sentences scaled by all added topic sentences (sentences with topic that does not occur in the earnings announcement) (standardized)
Sent_Joint_Topics	Difference between positive and negative sentences scaled by all joint topic sentences (sentences with topics that occur in both earnings announcement and article) (standardized)
Sent_EA	Difference between positive and negative words in the earnings announcement scaled by the sum of positive and negative words (standardized)
Sent_Art	Difference between positive and negative words in the article scaled by the sum of positive and negative words

Notes: This table shows variable definitions. All continuous variables (except for MABHR, production, and tone variables) are winsorized at the 1% level due to outliers. Continuous production and sentiment variables are standardized to facilitate the interpretation of coefficients.

Appendix 3.A2: Statement topic

Panel A: Topic descriptives

TextType	N with topic	mean	sd	min	p25	p75	max
Articles	6,181	0.32	0.20	0.00	0.16	0.48	0.93
EAs	998	0.01	0.02	0.00	0.00	0.00	0.31

Panel B: Example paragraphs

Topic proportion	Paragraph
Carnival (0.850)	"I believe we would have recovered had it not been for these two unfortunate incidents," Mr. Arison said in a telephone interview Tuesday. Mr. Arison said it was his decision to step down as CEO and that he had been discussing succession plans with the board for years, adding that the move likely had been delayed by the recent mishaps. Stuart Subotnick, Carnival's lead independent director, confirmed that Mr. Arison wasn't pressured to step down as CEO. "It was not the board pushing Micky [out]," Mr. Subotnick said in a separate interview. "It was the other way around."
Juniper (0.841)	Telecom companies, such as Verizon Wireless and AT&T Inc. are keeping spending flat for the second half of the year, having traditionally spent more after June, said Juniper Chief Executive Kevin Johnson. "We see similar trends with the top 15 service providers globally," Mr. Johnson said on a conference call with analysts. Any time a company misses like this, they do have to rebuild their credibility, said Simona Jankowski, a Goldman Sachs analyst who Wednesday removed Juniper from a group of favored stocks called its "conviction" list. "They will be in the penalty box, for at least a quarter."
U.S. Steel (0.825)	Sam Halpert, who manages the stake at Van Eck, said he recently proposed to U.S. Steel that it divide itself into three units overseen by a holding company, forcing each division to be more transparent and turn a profit while exploiting some tax benefits. Soon after, Mr. Surma and U.S. Steel's finance chief, Gretchen Haggerty, called Mr. Halpert. "They said they had people on the board looking at all kinds of different possibilities," Mr. Halpert recalled.

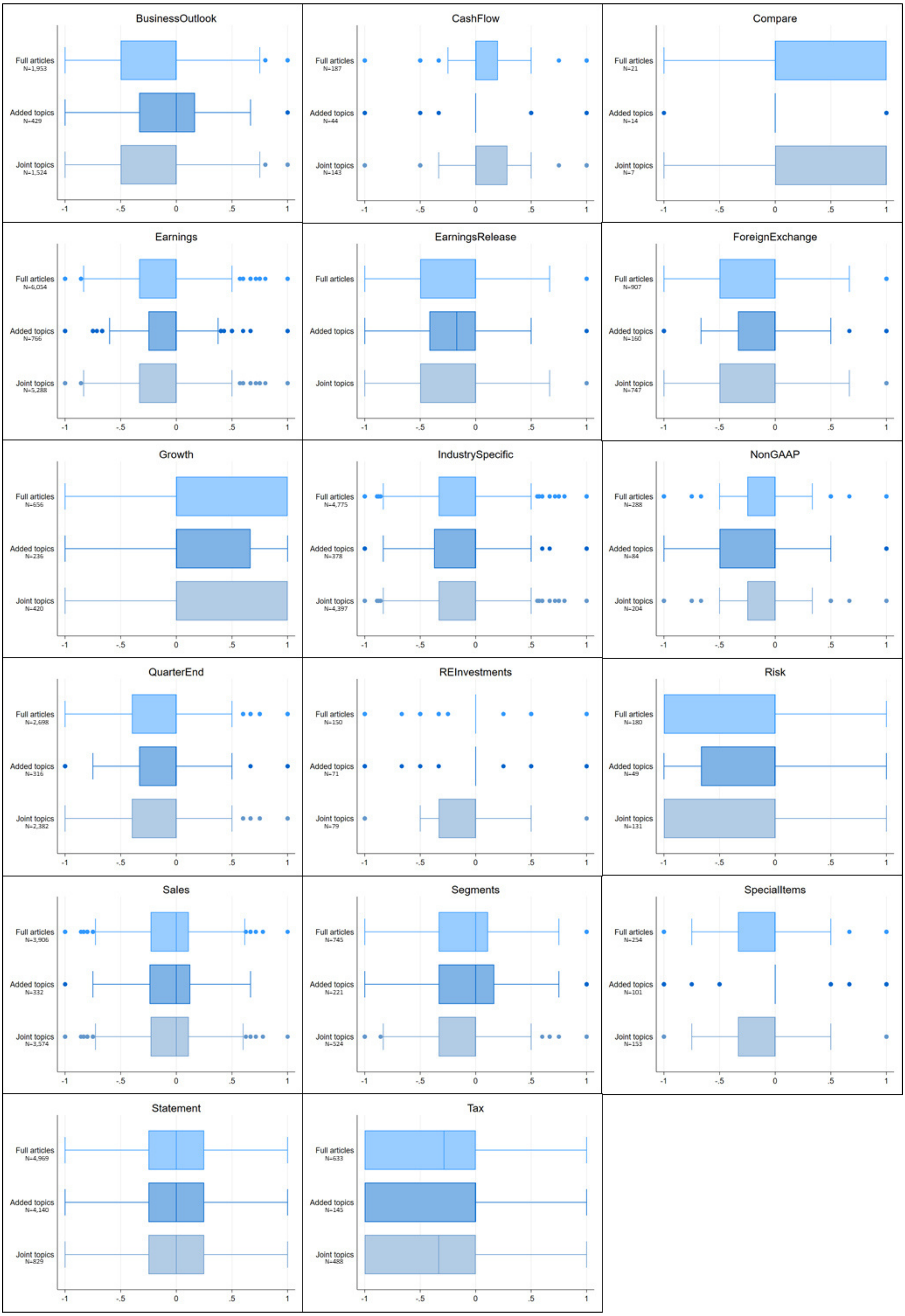
Notes: This table shows statement topic descriptives (Panel A) and example paragraphs featuring this topic (Panel B).

Appendix 3.A3: Additional analysis (*Emphasis)* [Table for Figure 3.3]

Topic	mean	t-stat	sd	min	median	max	count
BusinessOutlook	0.044	12.90***	0.143	-0.507	0.039	0.816	1,731
CashFlow	-0.069	-9.58***	0.118	-0.422	-0.052	0.277	270
Compare	-0.053	-2.54**	0.055	-0.151	-0.048	0.032	7
Earnings	0.244	77.10***	0.233	-0.825	0.267	0.863	5,405
ForeignExchange	0.010	3.61***	0.093	-0.358	0.007	0.779	1,146
Growth	0.063	16.14***	0.142	-0.597	0.062	0.598	1,323
IndustrySpecific	0.023	8.12***	0.193	-0.749	0.018	0.834	4,560
M&A	0.038	12.65***	0.088	-0.384	0.027	0.514	869
NonGAAP	-0.097	-14.12***	0.128	-0.570	-0.077	0.312	345
RealEstateInvestments	-0.012	-0.95	0.184	-0.661	0.016	0.532	208
Risk	-0.024	-2.95***	0.151	-0.438	-0.008	0.628	343
Sales	0.069	28.05***	0.147	-0.708	0.044	0.907	3,576
Segments	-0.024	-5.52***	0.112	-0.533	-0.004	0.348	674
SpecialItems	-0.038	-6.86***	0.099	-0.400	-0.017	0.217	328
Tax	-0.001	-0.25	0.098	-0.433	0.007	0.355	696

Notes: This table provides information about the additional analysis (*Emphasis*) in financial press articles. *Emphasis* excludes the topics “EarningsRelease”, “QuarterEnd”, and “Statement”. ***, **, * indicate that the mean topic loading is statistically significant different from zero the 1%, 5%, and 10% level, respectively based on a t-test.

Appendix 3.A4: Sentiment distribution



Appendix 3.A4 shows the box plots for the sentiment of the various topics on the sentence level.

4 Corporate social responsibility and stakeholder attention

Ann-Kristin Großkopf[26]

Working Paper[27]

Abstract: While there is ample evidence that investors use non-financial disclosure, evidence for non-investor stakeholders is rare. I use a sample of Fortune's "100 best companies to work for" list firms between 2005 and 2017 to provide evidence on stakeholders' attention following positive, externally validated CSR disclosure. Measuring the stakeholder attention using abnormal Google search volume, I find a significant increase in attention following the list publication. This increase is more pronounced for private firms, and the top 10 list ranks. I find that the increase in attention is not followed by an increase in applications to the list firms. In contrast, firms seem to increase their CSR activities in employee-related matters, as they pay less non-financial misconduct fines after being featured on the list. Overall, I find that non-investor stakeholders pay attention to non-financial information but fail to act upon it, whereas firms change their CSR activities when anticipating increased stakeholder attention.

JEL Classification: 14, G14, J28

Keywords: Corporate social responsibility (CSR), stakeholder reactions, financial media, employment

Acknowledgments: For valuable comments and suggestions, I thank my dissertation adviser Jörg-Markus Hitz, Nico Lehmann, Thorsten Sellhorn, Katharina Weiß, and workshop participants at LMU Munich.

---

[26] Georg-August Universität Göttingen, Faculty of Economic Sciences, Chair of Accounting and Auditing, Platz der Göttinger Sieben 3, 37073 Göttingen, Germany.

[27] First version of the working paper: September 2019. This version: February 2020.

4.1 Introduction

Regulators such as the EU frequently state that corporate social responsibility (CSR) disclosures should not only meet the information demand of investors but also of other stakeholders like consumers or employees (e.g., Directive 2014/95/EU, recital 3). However, there is little evidence whether and through which channels stakeholders use non-financial information and how they react to its disclosure. This lack of evidence is also due to the fact that stakeholder information demand changes over time and is difficult to observe because it is often unrelated to monetary incentives (Christensen et al. (2019); Hombach and Sellhorn (2018)). Yet, it is essential to understand how non-investor stakeholders use non-financial information, for example, to draft effective legislation. In this paper, I aim to provide evidence on the link between stakeholders and non-financial disclosure. To that end, I examine two research questions. First, I examine whether stakeholders increase their attention to a firm after the release of externally validated non-financial information. Only if stakeholders pay attention to the information, it is likely to observe any reaction on their behalf. Second, I investigate how this attention might have long-term consequences for firms based on stakeholder reactions (e.g., in the form of job applications) and their own (pre-emptive) actions (e.g., their CSR investments).

I examine the Fortune's "100 best companies to work for" list (BC list) between 2005 and 2017 as a specific form of non-financial disclosure. The BC list ranks firms according to an employee and a culture audit survey and covers areas such as respect, fairness, or benefit programs. As the annual ranking is disseminated trough a well-known information intermediary, it is more likely that stakeholders will become aware of this information compared to other channels through which firms publish their non-financial disclosure. Additionally, the publication in a newspaper makes the information more credible than disclosure by the firm itself (Kothari et al. (2009)). Further, the ranking is easy to understand without further background knowledge and thus, facilitates the comparison of various firms. A key point of the setting is that the list information is disseminated on two different dates. This dissemination facilitates the distinction between stakeholder and investor reactions. I use the first disclosure of the BC list as event 1 and the official issue date as event 2. Another important feature is that the list contains information about a specific aspect of non-financial information, namely employee-related CSR activities. This information is externally validated and should be most valuable to (potential) employees of the Fortune's "100 best companies to work for" list firms (list firms). Thus, the setting allows me to isolate the reaction of a specific stakeholder group.

For the stakeholders, the BC list is a positive signal about the CSR performance of the firms, which they consume through an information intermediary. It extends the information set of stakeholders and reduces the uncertainty about the working conditions in a list firm. Due to this information content, I expect an increase in stakeholder attention following the publication of the BC list. As the information is valuable for stakeholders when it enters their information set, regardless of the exact timing, I expect an increase in stakeholder attention when the information is first disclosed (event 1) and when it is more widely disseminated (event 2).

Based on the new information, the stakeholders update the expected net utility they might obtain from working for a list firm. If the updated utility based on the BC list exceeds the utility from their current employment, this might trigger stakeholder reactions in the form of job applications to the BC firms to maximize the utility they obtain from their employment (Hombach and Sellhorn (2018)). Consequently, I expect an increase in applications to list firms after they are featured on the list.

Additionally, the BC list facilitates the monitoring of employee-related CSR for the stakeholders. Consequently, list firms are likely to further invest in their CSR activities to remain on the BC list and signal their CSR strength to (potential) employees and prevent adverse stakeholder reactions. Therefore, I expect an increase in CSR performance for the list firms.

To answer my first research question regarding the stakeholder attention, I examine the stakeholder attention following event 1 and event 2. I measure the stakeholder attention using the abnormal Google search volume index (ASVI). In contrast to prior literature, which uses ASVI based on ticker to capture (retail) investor interest, I follow Madsen and Niessner (2019) and calculate the ASVI based on name searches to capture stakeholder attention. Therefore, my proxy for stakeholder attention is independent of channels such as the media but is initiated by the stakeholders themselves.

In line with my expectations, the results for the ASVI tests show positive and statistically meaningful results (at the 1% level) following event 1 and event 2 and a generally increased level of attention. Stakeholders pay attention to both ranking improvements and deteriorations, while the difference between those two groups is not statistically significant. The increase in attention is stronger for firms with a very good ranking (e.g., among the top 10/20 firms). Overall, my results suggest that stakeholders pay attention to the publication of non-financial information.

However, the name-based ASVI might also capture the interest of retail investors, who use the name of a firm to look for financial information online instead of its ticker. To alleviate these concerns, I compare the short-term stakeholder reaction with the market reaction following the BC list publication. To that end, I estimate three-day market-adjusted abnormal returns (MAR) for both events.

The BC provides valuable information content for the investors because employees with a high level of satisfaction are likely to improve their output and remain at their current employer, which reduces the uncertainty of the future cash flow. For investors, the information is most valuable when it is new. Therefore, I expect a significantly positive market reaction to the first event. I do not expect to find a significant market reaction for the second event, because the wider dissemination of the BC list information in Fortune does not provide new information content. In line with my expectations, I find a positive abnormal return following event 1, which is significantly different from zero and non-event returns at the 1% level. In contrast, I find no significant results following the second event. When splitting the sample into strictly positive and negative news, the analyses show that the BC list only provides valuable information content for the positive information sample. These results suggest a difference in information processing between investors and stakeholders, which supports the construct validity of my proxy.

To further alleviate concerns about the construct validity of the ASVI, I offer three additional identification steps based on different list characteristics and Google search terms. First, I exploit that the BC list features both private and public firms. As the private firms have fewer investors, an increase in ASVI is less likely attributable to (retail) investor attention and more likely based on stakeholders. I find that the abnormal ASVI for private firms is nearly twice as high following both events and significant at the 1% level.

Second, I compare name-based and ticker-based searches. Prior literature uses this proxy to identify investor attention (e.g., Da et al. (2011)). Consequently, and aligned with the results for the market reaction, I expect that the ticker ASVI is only significantly different from zero for event 1. Yet, I find significant results for both events. To rule out that high-noise tickers bias the results, I eliminate all tickers with a noise level above 85% based on the data provided in deHaan et al. (2019) from the sample. I find that the ASVI decreases in magnitude, whereas the significance levels for the events remain mainly unchanged. Nevertheless, the ASVI magnitude for name searches for both listed and private firms largely exceeds the results I find for ticker searches.

In the last step, I compare ticker- and name-based searches following the BC list publication with the search activity after the announcement of quarterly earnings. As earnings announcements are among the most important events for investors, the results provide evidence which search term investors are most likely to use. I find that, though the ASVI is positive and highly statistically significant for both search terms, the ticker ASVI is nearly twice as high as the name-based ASVI. This magnitude suggests that investors are more likely to use tickers than name searches. Based on all identification steps, I am confident that the name-based ASVI captures stakeholder attention. Cross-sectional analyses show that stakeholder attention is more pronounced for firms with more negative press coverage, while it increases when the public image is negative.

To answer the second research question regarding long-term stakeholder and firm reactions, I use two different analyses. First, I examine the stakeholder reaction following their increased attention after the list publication. To that end, I focus on labor market reactions, as the information of the BC list is closely linked to this stakeholder group. To alleviate concerns that the increase in stakeholder attention might emanate from other stakeholders such as customers instead of employees, I analyze the ASVI based on an employee-specific search term. I collect Google Trends data based on the name of the BC firms and add the term “jobs” to the search request. Though the SVI for this search term is generally lower, I find a significant increase in ASVI for both private and listed firms after both events (1% level).

Next, I analyze the stakeholder reaction by examining the development of job applications to a firm after being featured on the list for the first time. I find that there is an increase in applications but that this increase is not statistically significant neither for listed nor private firms. Overall, it seems like stakeholders pay attention to the publication of stakeholder related non-financial information but fail to act upon it.

For my second analysis, I examine the long-term CSR investment decisions of the list firms. I proxy these CSR investments using the non-financial corporate misconduct fines a firm must pay due to employee-related topics (e.g., due to workplace safety issues). If firms invest in their employee-related CSR performance, this should decrease the number of violations discovered by federal agencies or the number of lawsuits filed against the firms. I estimate a differences-in-differences (DiD) model using the first-time list inclusion as treatment and a propensity score matching (PSM) matched sample based on all non-list Compustat firms as controls to assess if list firms increase their CSR activities after being featured on the list for the first time. In line with my expectations, I find a significantly negative average treatment effect for the list firms. Overall, the results imply that firms anticipate the increase in stakeholder attention and increase their CSR activities to prevent adverse stakeholder reactions.

My paper contributes to the scarce literature about how stakeholders perceive non-financial information and how they react to its release. It is essential that we understand which non-financial information meets the information demand of stakeholders and how they react to its disclosure to draft effective non-financial reporting legislation, which meets the aims of regulators such as the EU. My paper builds on Madsen and Rodgers (2015), who examine stakeholder attention to corporate disaster relief, e.g., for hurricane victims and the resulting financial benefits for the contributing firms. While their proxy for stakeholder attention relies on newspaper coverage, which is subject to the coverage decision of journalists, the Google searches used in my paper are stakeholder initiated and, therefore, directly capture their interest. Additionally, employee-related CSR activities are likely part of the internal long-term CSR strategy of a firm, and thus, my paper captures another CSR dimension than Madsen and Rodgers (2015), which further enhances our understanding of the link between stakeholders and CSR.

As the various stakeholder groups such as customers or employees are rather diverse, it is difficult to assess how they process the non-financial information and how it might trigger stakeholder reactions. Building on the experimental design of Greening and Turban (2000), I isolate the reaction of a particular stakeholder group, i.e., current and future employees and examine if they react to the publication of non-financial information, e.g., by pursuing jobs from BC list firms. Additionally, I provide evidence on how increased stakeholder attention might impact the CSR investment decisions of a firm and therefore extend prior literature by capturing the response of firms to increased stakeholder attention.

The paper proceeds as follows. Section 4.2 provides insights into the information content of the BC list, reviews the related literature, and develops predictions. Section 4.3 outlines the research design choices and results for the stakeholder attention analyses. Section 4.4

presents the research design and results regarding the labor market reaction and investment effect analyses, while Section 4.5 concludes.

4.2 Background and related literature

In this section, I present background information on my setting, review the related literature, and develop my empirical predictions.

4.2.1 Information content of the "100 best companies to work for" list

Since 1998, Fortune magazine publishes a list of the "100 best companies to work for" (BC list) on an annual basis. To create the list, Fortune collaborates with the *Great Place to Work Institute*. Robert Levering and Milton Moskowitz originally developed the list in 1984. The news outlet itself is not part of the BC list creation. The external commission of the list limits concerns raised by prior literature regarding sensationalism in the media to attract readers instead of focusing on accurate reporting (e.g., Ahern and Sosyura (2015)). As Kothari et al. (2009) state that the press is a more credible source than analysts or the firms themselves, the publication in Fortune should enhance the credibility of the list for investors and other stakeholders.

Originally, Fortune published the list in its January edition but moved it to the March issue from 2015 on. The print version of the magazine arrives at the newsstand up to three weeks prior to the official publication date. In more recent years, Fortune has additionally published the list online on a predetermined date before the release of the current issue.[28] To identify the exact event dates, I use the ProQuest database and look for the first mention of the list, e.g., via firm press releases (Faleye and Trahan (2011)). I use the first disclosure as the date when the information becomes publicly known (event 1). In my setting, this first disclosure is mainly linked to the upload of the Fortune BC web page. As the second event of interest, I use the issue date of Fortune magazine containing the BC list (event 2) because the information is more widely disseminated on that date. The time lag between the issue and the first disclosure is present throughout the sample period. On average, the Fortune issue date is 14 days after the first disclosure of the list results.[29]

[28] Past Fortune list web pages are available online from 2006 on. That is why I assume Fortune has started publishing the list online on a specific date in that year. The Great Place to Work Institute announces the next publication date on its web page, while Fortune provides the web page of the list. For 2018 and 2019 the announcement date matches to web page release of the new BC list. I tried to contact both Fortune and the Great Place to Work Institute to get insights into their timelines but did not get any feedback from either one of them.
[29] Table 4.3 provides the exact event dates for both events in each sample year.

The distinction between the two event dates provides a key feature of the setting, which I exploit to distinguish between investor and stakeholder attention.

To be considered for the list, firms must apply approximately eight months before its publication at the beginning of the year. The application is restricted to firms that have been operating in the USA for more than seven years, with more than 1,000 employees. Approximately 400 private and public firms apply for the list each year (Guiso et al. (2015); Gartenberg et al. (2019); Gartenberg and Serafeim (2019)).

The list ranks the firms according to a score, which is the result of two aggregated scores: The Trust Index© employee survey and the Culture Audit Survey. For the ranking, it is of no interest if the firm is listed or private, as the score only determines the 100 best firms. Two-thirds of the final score derive from the Trust Index© employee survey, which captures the employees' assessment of the firm. The survey consists of 57 questions and covers areas such as credibility, respect, fairness, pride, and camaraderie. The *Great Place to Work Institute* sends the survey to at least 400 random employees from all executive levels. The response rate for the survey is about 60%. The applying firms are not allowed to interfere with the responses other than to encourage employees to respond to the survey. The employees send their replies directly to the institute. The remaining part of the score derives from a culture audit, for which the management of the firm has to answer questions about the demographic makeup, pay and benefits programs, and culture of the firm (Garrett et al. (2014); Edmans (2011); Guiso et al. (2015); Gartenberg et al. (2019)). Therefore, the final ranking is an external assessment of the CSR activities of the firms, which is then disclosed to the stakeholders. In contrast to voluntary CSR reports published by the firms, the specific questions prevent firms from covering up or leaving out areas in which they do not perform as well as their competitors. The external assessment supports objectivity and therefore promotes credibility.

4.2.2 Prior literature

Overall, my paper contributes to the scarce literature about the link between stakeholder theory and CSR. It relates to CSR disclosure, as the BC list features an external assessment of employee-related CSR activities disseminated to the stakeholders through an information intermediary. In this setting, the non-investor stakeholders are the key target group of the information. Additionally, it relates to literature about CSR activities, as I examine the non-financial activities of firms in anticipation of increased stakeholder attention.

It is reasonable to assume that stakeholders are interested in non-financial activities that directly affect them. For example, employees likely care about issues such as workplace safety. To that end, Chen et al. (2018) and Christensen et al. (2017) show that firms improve the safety of workspaces (i.e., adapt their CSR activities) when they have to initiate mandatory safety disclosure or disclose the information more visibly. Fiechter et al. (2019) show that firms increase their CSR activities prior to the effective date of the EU disclosure mandate to prevent adverse stakeholder reactions due to simplified monitoring by stakeholders.

The literature about non-financial disclosure mainly focuses on investor reactions and their general use of this disclosure type (e.g., Dhaliwal et al. (2011); Flammer (2013); Krüger (2015)). Focusing on the BC list, Edmans (2011) and Edmans (2012) examine long-term market effects and find that list firm portfolios earn long-term excess stock returns. Faleye and Trahan (2011) examine first-time list inclusions between 1997 and 2005 and find significantly positive short-term returns. These results suggest a positive association between the BC list inclusion and financial performance.

The literature about stakeholder attention to CSR disclosure and their use of the published information is limited. However, some papers show that stakeholder interest in non-financial aspects goes beyond the information that directly affects themselves. Focusing on employees as an important stakeholder group, Greening and Turban (2000) use an experimental design to show that future employees are more likely to apply to firms that behave socially responsibly. Bunderson and Thompson (2009) find that employees that are dedicated to their work and find a kind of self-fulfillment or broader meaning in their work are willing to work for lower wages, i.e., when the firm engages in socially valuable behavior. To that end, Guiso et al. (2015), Garrett et al. (2014), Gartenberg et al. (2019), and Gartenberg and Serafeim (2019) use the full dataset of successful and unsuccessful BC list applications and examine the link between corporate culture, employee trust, and employee purpose. The results imply that employees' belief in purpose is linked to better financial performance and disclosure quality. Additionally, this belief in purpose seems to vary with the ownership structure of a firm.

On a broader level, Madsen and Rodgers (2015) examine stakeholder attention to corporate disaster relief efforts. The results suggest that the stakeholder attention to corporate disaster relief programs ultimately results in financial benefits for the firms. While my paper builds on Madsen and Rodgers (2015), I differ from their work in three essential ways. First, I measure stakeholder attention using Google searches rather than newspaper coverage, as Cahan et al. (2015) find that firms actively manage their media image through their CSR activities. Additionally, Google searches are stakeholder initiated, while media reports are subject to the coverage decision of journalists. Second, I focus on CSR information that mainly affects (potential) employees and am thus able to isolate the reaction of a specific stakeholder group. Third, I examine CSR activities that target an internal stakeholder group and are likely part of a long-term CSR strategy rather than the reaction to an unexpected event.

4.2.3 Empirical predictions

4.2.3.1 Stakeholder attention

Comparable to Servaes and Tamayo (2013), who argue that consumers need to be aware of CSR activities to affect firm value, stakeholders need to be aware of the superior employee-related CSR performance of a firm to react to it. To that end, Christensen et al. (2017) show that awareness regarding CSR information depends on where firms disclose information. In the case of the BC list, the stakeholders consume the information through the press as an information intermediary, which likely renders the information more credible (Kothari et al. (2009)). This credibility is further enhanced, as the press does not directly produce the list but commissions it from an independent institute. The publication in Fortune magazine makes the information easy-to-access, as it is widely disseminated, and the standardized ranking simplifies the interpretation of the information. It is more likely that stakeholders such as employees access this kind of information through the press instead of using traditional capital markets disclosure or CSR reports (Hombach and Sellhorn (2018)).

The information contained in the BC list reflects the firms' CSR performance regarding their employees. Contrary to the EU, there is no mandate for US firms to publish non-financial disclosure. Consequently, stakeholders in this setting might not be able to obtain the information of the BC list from other sources. For stakeholders such as employees, the positive information conveyed in the BC list reduces the uncertainty about the working conditions in a specific firm and additionally allows them to compare their current employer to listed ones. As the BC list presented to the stakeholders conveys valuable information content for them, I expect an increase in their online search activity for the firm.

The information content of the BC list should be valuable for stakeholders when it enters their information set, independent of the exact timing. The list might enter the information set at the first mention of the list, e.g., through a press release (event 1) or when the Fortune issue is published (event 2), and the information is likely to reach a broader audience. Overall, I expect a significant increase in stakeholder attention, proxied by online search activities, for both events.

The degree of employee satisfaction conveyed in the BC list is also beneficial for shareholders because motivated employees are more likely to make for an effort at work, which improves their output. Additionally, good working conditions facilitate staff retention, as satisfied employees are less likely to leave the firm. This retention is especially important in industries in which highly qualified employees are scarce (Edmans (2011)). Overall, the inclusion in the BC list reduces the uncertainty of the future cash flow for investors. Consequently, I expect a significant and positive market reaction to the publication of the BC list. In contrast to stakeholders, the information about the list is most valuable to investors when it is new (event 1), as investors are more likely to price the information directly. The wider dissemination of the BC list (event 2) does not provide any new information content. Therefore, I expect a significant market reaction for the first event but not for the second one.

4.2.3.2 Determinants of stakeholder attention

Additionally, I examine which firm characteristics shape stakeholder attention. My main variables of interest for this cross-sectional analysis relate to the public familiarity and the image of a firm among stakeholders. I proxy the public familiarity using the amount of press coverage of a firm. If firms are featured in the press more often, the stakeholders might be familiar with the firms, and thus the information that enters their information set might be stale and their search activity lower compared to less publicly familiar firms. (Hombach and Sellhorn (2018)). Therefore, I expect a negative association between the amount of prior press coverage and the abnormal search volume. I proxy the public image of a firm using the sentiment of the firms' press coverage. I do not have a signed prediction for this association. On the one hand, stakeholders might pay less attention to positive non-financial information if the firm is continuously featured negatively in the press as the signal might not be strong enough to change their prior beliefs about a firm. On the other hand, stakeholders might be more interested in the firm if the new information does not match their previous opinion.

4.2.3.3 Labor market reaction and investment effects

In addition to the short-term stakeholder and investor effects, being featured on the list might have two different long-term effect channels for the list firms. The first channel, which captures the labor market reaction, might impact the access of a firm to high-quality employees and is directly tied to stakeholder attention. As outlined above, I expect that the BC list attracts the attention of stakeholders, especially of (potential) employees, as the information content is linked to employee-related CSR. Based on the list information, stakeholders update the expected net utility they might obtain when working for a list firm and compare it to their current employment. The additional information the stakeholders obtain through their Google searches might further support their belief update.

Greening and Turban (2000) argue that applicants are more likely to pursue a job when a firm acts socially responsible, i.e., the expected net utility for list firms signaling good CSR performance might be higher compared to the current employer. Therefore, the positive CSR signal (list inclusion) might prompt stakeholder reactions in the form of job applications to the firm to maximize stakeholders' utility from their employment. (Hombach and Sellhorn (2018)). Applying to a list firm after the publication of the BC list would be comparable to an investor who buys shares of a firm after learning about positive financial news. The belief update is supposedly larger when Fortune features the firm on the list for the first time, as the expected net utility before the BC list publication is not affected by prior list information. Overall, I expect a significant increase in applications to list firms once they are featured on the list for the first time.

Second, the list inclusion might impact the long-term CSR investment decisions of a firm. These investment decisions might be a result of the anticipated increase in stakeholder attention. When applying for the list for the first time, the firms lack a benchmark of how well they are doing compared to their competitors. Consequently, they might be ranked among the top 100 immediately, or they might be rejected. If a firm does not make the list, it can use the rejection as an incentive to further invest in employee-related CSR activities to be successful when applying in the next year. If firms invest in their employee-related CSR performance and start new activities, they might invest in essential factors like workplace safety or implement anti-discrimination initiatives, for which further investments should decrease the number of violations discovered by federal agencies or the number of lawsuits filed against them.

After being featured on the list for the first time, the BC list creates a monitoring opportunity for the stakeholders, as it is easy to compare the firms and benchmark their employee treatment with other list firms (Christensen et al. (2019)). Therefore, the list firms are likely to further invest their high standard of employee treatment and their level of CSR activities to signal their CSR strengths to potential employees and prevent adverse stakeholder reactions (Zyglidopoulos et al. (2012); Fiechter et al. (2019)). These investments should further reduce the likelihood of federal conduct violations or legal actions against the firm. Consequently, I expect to observe fewer employee-related corporate misconduct fines once the firms are featured on the BC list. Overall, I expect a negative association between the amount of non-financial corporate misconduct fines and the first-time list inclusion.

4.3 Effects of CSR disclosure on stakeholder attention

In this section, I present and discuss the key aspects of the research design for the stakeholder attention analyses: the sample composition and data sources, the main test of the stakeholder attention, the test of the investor reaction, and the supporting identification steps.

4.3.1 Sample and data

The Fortune lists and the corresponding firm identifiers are available at Alex Edmans' personal web page up to 2012. I continue his firm-identifier list up to 2017 based on the Fortune BC list articles. I collect market data from CRSP and firm fundamentals from Compustat. Analyst data is available via IBES. I obtain newspaper related data from RavenPack News Analytics (Barber and Odean (2008); Drake et al. (2014)). To test the increase in stakeholder attention following the list inclusion of a firm, I obtain search volume index (SVI) data from Google Trends. Appendix A gives detailed explanations for the collection of Google Trends data and provides some general information about this dataset.

Table 4.1: Sample description

Panel A: Sample selection stakeholder attention		
Selection criteria	Unique firms	Firm-years
Start: Top100 firms between 2005 and 2017	257	1300
Less observations of firms:		
private firms	(139)	(721)
with confounding events (-2/+2)	(12)	(30)
with missing data	(4)	(103)
Sample stakeholder attention	**102**	**446**
Panel B: Sample selection matched sample		
First-time firms	45	270
Less observations of firms:		
missing pre or post	(7)	(42)
firms without match	(4)	(24)
missing data	(0)	(24)
Sample before matching	34	180
Matched Sample	**68**	**377**
Panel C: Sample selection private firms		
Start: Top100 firms between 2005 and 2017	257	1300
Less observations of firms:		
listed firms	(118)	(579)
with missing (Google) data	(16)	(40)
Sample stakeholder attention (private firms)	**123**	**619**

Panel D: Sample distribution per year				
Year	Stakeholder attention		Matched Sample	
	N listed firms	N private firms	N Treatment	N Control
2004	0	0	7	7
2005	45	24	10	10
2006	41	34	15	15
2007	40	45	19	20
2008	36	50	17	20
2009	32	46	17	22
2010	29	50	19	20
2011	33	50	16	17
2012	31	48	13	14
2013	33	50	11	12
2014	30	52	9	12
2015	32	56	8	9
2016	30	59	7	7
2017	34	55	7	7
2018	0	0	5	5
Total	*446*	*619*	*180*	*197*

Notes: This table provides details on the sample selection process for the stakeholder attention sample for listed firms (Panel A) and the matched sample (Panel B). Panel C shows the sample selection process for private firms. Panel D provides information on the annual distribution of each sample outlined in Panel A, B, and C.

As Google Trends data is not available until 2004, and because I require ten weeks of SVI data before the list publication, my sample consists of list firms between 2005 and 2017. The list features both private and listed firms. I move 721 private firm observations from the main sample to the private firm sample. Next, I exclude all firms with earnings announcements within a five-day window around the publication of the Fortune list, as this kind of information likely affects public interest in the firm. Finally, I exclude 103 observations with missing data. This leads to a final stakeholder attention sample of 446 firm-year observations, which are based on 102 firms. Panel A to C of Table 4.1 outline the sample selection process. Panel D of Table 4.1 further shows the sample distributions per year.

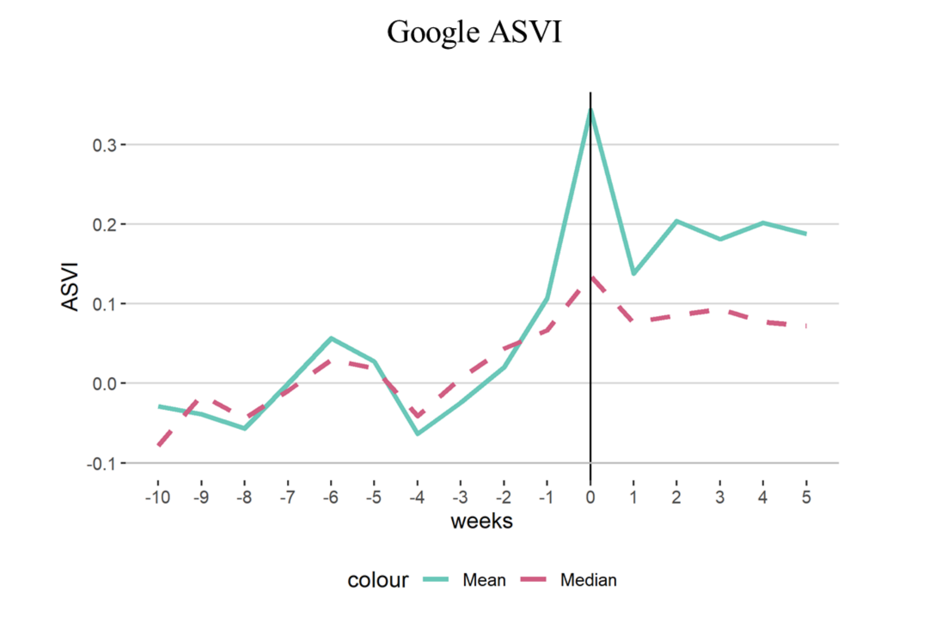
4.3.2 Test of stakeholder attention

While prior literature mainly uses the ticker-based abnormal search volume index (ASVI) to capture investor attention (e.g., Drake et al. (2012); Da et al. (2011); Drake et al. (2017)), I follow Madsen and Niessner (2019) and use name-based searches to capture stakeholder attention. I examine the ASVI on the day of the list publication (event 1) and the Fortune issue date (event 2). As outlined in the predictions, I expect a significant increase in stakeholder attention following both events.

Following deHaan et al. (2019) and Drake et al. (2012), I measure ASVI as the SVI for firm i on the day of the list publication (event 1) less the mean SVI for the firm on the same weekday over 10 weeks prior to the list publication, scaled by the same 10-week mean. As I want to compare the stakeholders' attention after the list information becomes public to the level before it is publicly known, I use the same 10-week period to calculate the ASVI for event 2 based on the Fortune issue date weekday.[30]

Figure 4.1 shows the mean and median stakeholder attention following the BC list publication proxied by the ASVI based on firm name searches. All firms in this sample are listed firms. The vertical line denotes event 1.[31] Figure 4.C1 (Appendix) shows the results for the absolute SVI.

[30] The results of the stakeholder attention analysis remain mainly unchanged when using the same SVI mean (based on the weekday of event 1) to calculate the ASVI for both events.
[31] To facilitate the comparison, the ASVI in this figure is based on the weekday of event 1 for all weeks.



week	mean	t-stat	week	mean	t-stat	week	mean	t-stat	week	mean	t-stat
-10	-0.029	-1.123	-6	0.057	2.334**	-2	0.020	0.820	2	0.204	5.274***
-9	-0.039	-1.623	-5	0.028	1.175	-1	0.107	3.3901**	3	0.181	5.397***
-8	-0.057	-2.230**	-4	-0.063	-2.604**	0	0.343	6.073***	4	0.202	5.717***
-7	0.000	-0.010	-3	-0.024	-0.952	1	0.138	3.868***	5	0.188	4.879***

Figure 4.1: Cross-sectional average ASVI around Fortune list publication
This figure shows the cross-sectional mean and median Google abnormal search volume index (ASVI) around the week of the list publication based on the weekday of event 1. Week 0 (vertical line) is the week of the list publication based on the first mention of the publication. The sample is based on list publications between 2005 and 2017 for listed firms (N=446). ***, **, * indicate statistical significance at the 1%, 5%, and 10% level, respectively (H_0=0).

The figure shows an increase in stakeholder attention leading up to event 1 (week 0). This increase is statistically significant at the 1% level. The ASVI remains positive and statistically significant in the following weeks. The AVSI decreases after the initial surge in attention but seems to increase after the issue publication (approximately two weeks later). Afterward, the ASVI remains comparably stable and significantly different from zero. This result suggests that the BC list increases stakeholder attention on a long-term basis, as the information might enter the information set of stakeholders at any point in time.

Table 4.2 presents the cross-sectional means and t-tests for the ASVI for both event dates (Panel A and C) and a comparison of the ASVI based on various list characteristics (Panel B and D). While column (1) in Panel A corresponds to the results of Figure 4.1, the overall results in Panel C (column (1)) additionally show a significant increase in stakeholder attention for event 2. Though the mean for event 2 is lower, the results are still significantly different from zero on the 1% level.[32] Taken together, the results suggest that the stakeholder pay attention to both the first disclosure of positive CSR information and its wider dissemination via an information intermediary, which is in line with my expectations.

While a place among the 100 best companies is an overall positive signal for both stakeholders and investors, it can be moderated or reinforced based on the specific list rank of a firm. On the one hand, the signal might be stronger if a firm is featured on the list for the first time or when the rank of the firm improves. On the other hand, the signal might be weaker when the firm is still featured on the list, but lost ranks compared to the prior year. Consequently, I also examine whether the information content of the list differs for firms based on specific list characteristics.

Columns (2) to (5) in Table 4.2, Panel A and C show the results based on various characteristics such as a deteriorated or improved ranking for event 1 and event 2, respectively.[33] The results of all four sample splits remain highly significant (1% level) for the first event (Panel A). For the second event (Panel C), I find a highly significant ASVI for the sample splits based on an improved or deteriorated ranking (columns (2) and (3)), while the results for the top 10 ranks (column (4)) are not, and the first-time list inclusion (column (5)) only marginally statistically significant. Overall, the sample splits support my expectation that stakeholders pay attention to both events.

Table 4.2 Panel B (event 1) and D (event 2) present t-tests for the comparisons of list firms based on their ranking characteristics (e.g., top 10 firms vs. non-top 10 firms). I find a stronger increase in stakeholder attention for top list ranks, e.g., firms among the top 10 for event 1. The magnitude of this effect decreases with the list rank. In contrast, the difference between firms with an improved or deteriorated ranking is not statistically significant. This implies that stakeholders pay attention to the non-financial information regardless of the ranking development. For event 2, I do not find a significant difference in stakeholder attention based on list characteristics (Panel D). Taken together, the results suggest that stakeholder attention increases following the disclosure of non-financial information. This increase is independent of the concrete timing and the specific list rank,

[32] For event 2, I drop 12 firms due to confounding events around the issue publication.
[33] The improved ranking groups also features first-time firms.

which implies that the list inclusion itself provides valuable information content for the stakeholders.

Table 4.2: Stakeholder attention based on event dates

Panel A: Stakeholder attention event 1 (date first disclosure)					
	(1) Overall attention H_0=0	(2) Improved ranking H_0=0	(3) Deteriorated rank H_0=0	(4) Top 10 H_0=0	(5) First time H_0=0
N	446	261	172	37	45
mean	0.343	0.331	0.303	1.406	0.539
t-statistics	6.073***	4.869***	4.193***	2.834***	3.155***
Panel B: Stakeholder attention comparisons event 1					
	(1) Positive ranking (N=261)	(2) Deteriorated rank (N=172)	(3) Top 10 (N=37)	(4) Top 20 (N=71)	(5) First time (N=45)
Rank firms	0.331	0.303	1.406	0.900	0.539
Other firms	0.361	0.369	0.247	0.238	0.322
Difference	-0.030	-0.066	1.159	0.662	0.217
t-statistics	-0.258	-0.570	5.858***	4.371***	1.156
Panel C: Stakeholder attention event 2 (issue date)					
	(1) Overall attention H_0=0	(2) Improved ranking H_0=0	(3) Deteriorated rank H_0=0	(4) Top 10 H_0=0	(5) First time H_0=0
N	434	254	168	36	43
mean	0.201	0.230	0.148	0.211	0.328
t-statistics	4.602***	3.612***	2.658***	1.378	1.695*
Panel D: Stakeholder attention comparisons event 2					
	(1) Positive ranking (N=254)	(2) Deteriorated rank (N=168)	(3) Top 10 (N=36)	(4) Top 20 (N=70)	(5) First time (N=43)
Rank firms	0.230	0.148	0.211	0.187	0.328
Other firms	0.160	0.234	0.200	0.203	0.187
Difference	0.070	-0.086	0.011	-0.016	0.141
t-statistics	0.795	-0.964	0.071	-0.137	0.967

Notes: This table shows the ASVI for listed firms for event 1 (Panel A and B) and event 2 (Panel C and D) based on firm name searches. Panel A and C provide results of t-tests comparing the ASVI to zero. Column (1) provides information on the full sample (N=446), while columns (2) to (5) test different subsamples, e.g., for firms with an improved list ranking. Panel B and D provide results of t-tests comparing different list groups based on list characteristics. For example, column (1) compares the stake-holder attention of firms with an improved ranking (N=261) to the other list firms (with deteriorated or constant ranking; N=446-261=185). ***, **, * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

While I use the ASVI as a proxy for stakeholder interest in a firm, it is also possible that (retail) investors use name searches to look for the firm online. In this case, the ASVI would not only capture stakeholder attention but reflect both investor and stakeholder interest. To alleviate concerns about the construct validity of firm name searches to capture the stakeholder attention, I run several additional tests outlined below.

4.3.3 Test of investor reaction

To differentiate between investor and stakeholder attention, I exploit the timing of event 1 and event 2. As mentioned above, I find a significant increase in name-based ASVI for both events. However, I expect that investors only react to the first event, while there should be no significant market reaction for the second one because event 2 does not contain new information content for investors. This, in turn, would make it less likely that the increase in abnormal Google search volume might be attributable to (retail) investor searches, especially for the second event.

To assess the market reaction for the two events and to compare the results with those of the stakeholder analysis, I examine the three-day cumulative market-adjusted return (MAR) centered on the publication date of the list (event 1) and the Fortune issue date (event 2). As I focus on the US version of the BC list, I use the S&P500 index for the market-adjustment to proxy the market development without the publication of the list. I define *MAR* as the difference between the cumulative return of the BC list firms and the cumulative index firms. If the list is published on a non-trading day, I use the consecutive trading day as event date. Following Joos and Leung (2013) and Armstrong et al. (2010), I use three-day non-event market-adjusted returns as an alternative benchmark for the overall market reaction to the list publication. In line with the ASVI analysis, I exclude all firms that publish their earnings within a 5-day window (-2/+2) of the Fortune list publication to control for confounding events.

Table 4.3 reports the results of the overall market reaction for event 1 (Panel A) and event 2 (Panel B) using three-day MARs. Panel A shows an on average positive abnormal return across all 13 sample years, which is significantly different from zero at the 1% level. Additionally, I find that the event-day mean return is significantly different from the non-event mean return at the 1% level. Table 4.3 Panel B presents the same analysis for the Fortune issue date. The mean return for this event is not statistically different from zero. Further, the mean is not statistically different from the mean of non-event market-adjusted returns.[34] Overall, the market reactions suggest that investors value the information content of the BC list. In line with my expectations, the investors price the information when it becomes public (event 1) and do not react to its wider dissemination in Fortune magazine in a statistically meaningful way (event 2). These results contrast those of the ASVI analysis and make it less likely that the significant increase in ASVI following event 2 is based on investor searches.[35]

Though I generally expect the list inclusion to provide positive information content, the specific rank characteristics of firms might moderate and strengthen the investor reaction comparable to the stakeholder attention. For example, investors might regard an improved ranking as good news, while they consider a ranking deterioration to be bad news. This, in turn, would affect the results of my MAR analysis. Therefore, I split the sample into firms that are featured on the list for the first time or improve their rank (positive list ranking) and firms losing ranks compared to the prior years (negative list ranking).[36]

Table 4.C2 (Appendix) shows the results for the positive list ranking.[37] In line with the results for the full sample, I find a significantly positive MAR for the first event, while the MAR for the second event is not significantly different from zero. Table 4.C3 (Appendix) reports the results for the negative list ranking. In contrast to my prior results, the MAR is not significantly different from zero for neither event 1 nor event 2. These results suggest that while the first disclosure of an improved or first-time ranking provides valuable information content for investors, a deteriorated ranking does not. This finding diverges from the ASVI results because stakeholders seem to pay attention to both an improved and deteriorated rank as the increase for each subsample is statistically significant, and there is no significant difference between the two groups.

---

[34] As for the ASVI analysis, I drop 12 firms due to confounding events around the issue publication.

[35] To assess the robustness of the results, I re-estimate the market reaction for the first event using three-day cumulative abnormal returns (CAR) in line with MacKinlay (1997). The estimation window for the CARs comprises 175 days and ends 50 days before the annual publication of the list. Untabulated results show that the CAR is significantly different from zero. Thus, the results for the market reaction remain unchanged.

[36] I further test whether the market reactions yield significant results for the first list inclusion, the first two list years and top 10 or 20 ranks, respectively. I find no significant results for the first list inclusion (t-statistic 1.29), the first two list years (t-statistic 1.60), and top 10 ranks (t-statistic 1.33). I find a slightly positive market reaction for a top 20 ranking (t-statistic 1.75).

[37] I exclude firms with an unchanged ranking from this analysis.

Table 4.3: Overall Market Reaction event dates

Panel A: Market reaction event 1 (date first disclosure)

(1) Event Date	(2) List year	(3) N	(4) Cum.ret. (-1/+1)	(5) S&P500 (-1/+1)	(6) MAR (4)-(5)	(7) t-statistic (H_0= 0)
10jan2005	2005	45	-0.004	-0.004	0.000	0.042
09jan2006	2006	41	0.017	0.013	0.004	1.198
08jan2007	2007	40	-0.004	-0.004	0.001	0.337
22jan2008	2008	36	0.026	0.004	0.022	1.749*
22jan2009	2009	32	0.043	0.034	0.009	0.643
21jan2010	2010	29	-0.047	-0.052	0.005	1.103
20jan2011	2011	33	-0.013	-0.009	-0.004	-0.939
19jan2012	2012	31	0.027	0.017	0.011	1.861*
16jan2013	2013	33	0.014	0.007	0.007	1.821*
16jan2014	2014	30	0.008	0.000	0.008	1.634
05mar2015	2015	32	-0.017	-0.017	0.000	0.013
03mar2016	2016	30	0.012	0.011	0.002	0.324
09mar2017	2017	34	0.005	0.002	0.003	1.360
Mean		***446***	***0.006***	***0.000***	***0.005***	***2.945****
Non-event day mean return					0.001	
t-statistic (H_0 = non-event mean)					2.997***	

Panel B: Market reaction event 2 (issue date)

(1) Event Date	(2) List year	(3) N	(4) Cum.ret. (-1/+1)	(5) S&P500 (-1/+1)	(6) MAR (4)-(5)	(7) t-statistic (H_0= 0)
24jan2005	2005	43	-0.019	-0.006	-0.013	-2.496**
23jan2006	2006	39	-0.012	-0.014	0.002	0.443
22jan2007	2007	38	0.002	0.001	0.001	0.244
04feb2008	2008	35	-0.011	-0.030	0.019	1.233
02feb2009	2009	30	-0.013	-0.008	-0.006	-0.541
08feb2010	2010	28	0.013	0.007	0.006	1.621
07feb2011	2011	32	0.021	0.013	0.007	1.912*
06feb2012	2012	30	0.019	0.016	0.003	0.668
04feb2013	2013	33	0.017	0.009	0.008	1.023
03feb2014	2014	30	-0.016	-0.022	0.006	1.208
16mar2015	2015	32	0.012	0.004	0.008	1.985*
15mar2016	2016	30	-0.004	0.003	-0.007	-1.017
15mar2017	2017	34	0.008	0.003	0.004	2.167**
Mean		***434***	***0.001***	***-0.002***	***0.003***	***1.432***
Non-event day mean return					0.001	
t-statistic (H_0 = non-event mean)					-1.304	
t-statistic (event 1 = event 2)					1.279	

Notes: This table provides an analysis of the overall market reaction to the first disclosure of the BC list (Panel A) and the publication of the Fortune issue containing the BC list (Panel B). Column (4) reports the three-day cumulative return of sample firms centered on each event date. Column (5) presents the three-day cumulative return for the S&P500 index. MAR in column (6) is the difference between column (4) and column (5). I test whether the mean return for three-day MAR (column (6)) is significantly different from zero (H_0=0), which is denoted by the t-statistic in column (7). Non-event day mean return is the average of non-overlapping three-day market-adjusted returns for non-event days in all sample years. I test whether the mean return for three-day MAR (column (6)) is significantly different from the non-event day mean return (H_0=non-event mean return). Panel B additionally compares the results for event 1 and event 2. All variables are defined in Table 4.C5. ***, **, * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Collectively, the comparison of the stakeholder attention and the investor reaction implies that the information processing of stakeholders and investors differs regarding non-financial information, which allows me to isolate the stakeholder attention. This, in turn, supports the construct validity of the ASVI as a proxy for stakeholder attention.

4.3.4 Supporting identification steps

4.3.4.1 Design

I use three additional identification steps to alleviate further concerns that the name-based ASVI captures investor rather than stakeholder attention. All steps are based on different search terms and the corresponding ASVI. I run the supporting ASVI tests for both events for all search terms outlined in the supporting identification steps below.

First, I exploit the fact that the list consists of private and public firms. Public firms might have many investors, and consequently, many (potential) shareholders might look for the firm after the disclosure of positive news and affect the ASVI. In contrast, the more concentrated ownership of private firms reduces the number of investors looking for the firms online after the BC list publication and therefore minimizes a potential investor impact on the ASVI. If the name searches for public firms were mainly based on investor searches, I should find no significant increase in ASVI after the list publication for private firms for either event. In contrast, if I found a significant increase in name searches, comparable to my results for the public firms, this increase would most likely be based on stakeholder attention.

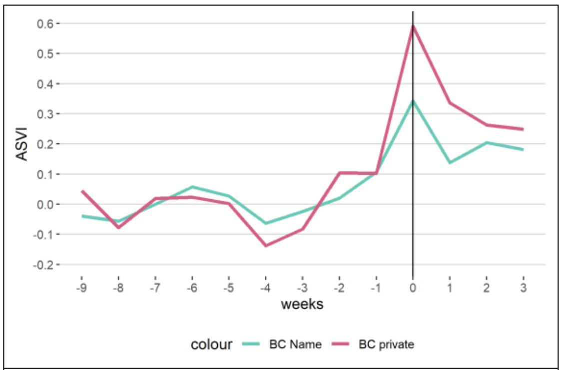
Second, I compare the name-based and ticker-based ASVI. Prior literature uses ticker searches as a proxy for investor attention (e.g., Da et al. (2011); Drake et al. (2012)). Based on my results for the market reaction, the ticker ASVI should only yield significant results for the first event. Significant results following the second event would indicate that less sophisticated investors look for firm information once they become aware of the BC list even though there is no new information content, which could also impact the results based on name searches. deHaan et al. (2019) show that many tickers are noisy due to alternative meanings of their letter combination (e.g., “cat”, which is the ticker for Caterpillar). Consequently, there are Google searches for tickers that do not intend to find any financial firm information but increase the SVI of the search term regardless. Based on a newly added Google Trends feature that allows for category searches (e.g., Finance-Investor), the authors develop a list of high-noise tickers. These tickers are likely to bias any results I might find for the ticker-based searches, as the term search frequency might vary completely unrelated to any information content. I exclude all firms from the sample whose ticker has more than 85% noise, according to deHaan et al. (2019).[38] I eliminate 48 observations due to high-noise tickers and repeat the analysis of the second step with the reduced sample.

Finally, I compare the magnitude of the ASVI for name and ticker searches after the BC list publication with a mainly financial event, namely the quarterly earnings announcement of a firm. As earnings announcements are among the most important investor events (e.g., Basu et al. (2013)), this analysis yields insights into which search terms investors are more likely to choose. Madsen and Niessner (2019) run a similar test to validate that their ticker-based proxy captures investor interest and find that increases in ticker searches after an earnings announcement are much larger than name search increases. If I were able to replicate their results with the BC sample, it would seem more likely that investors search with tickers after value relevant events, while name searches mainly capture stakeholder interest. For this comparison, I use the quarterly earnings announcements of list firms within a five-month (-3/+2) window of the list publication.

4.3.4.2 Results

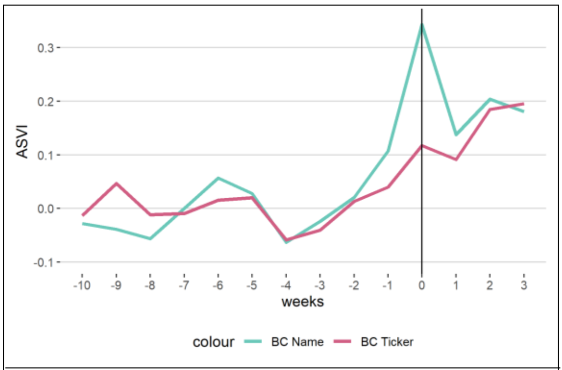
Figure 4.2 presents the results for the ASVI with respect to the supporting identification steps. The results for *BC Name* in all three panels correspond to the mean shown in Figure 4.1. As in Figure 4.1, the ASVI is calculated using the SVI 10-week mean for the weekday of event 1. *BC private* shows the ASVI for the private BC list firm (Panel A). *BC Ticker* refers to the same public firms as *BC Name* but uses their tickers for the online searches (Panel B). *RDQ Name and Ticker* refers to the firm name and ticker searches of list firms for the reporting date of quarterly earnings announcements of these firms (Panel C).

[38] At the moment, the deHaan et al. (2019) working paper only features noise levels for S&P500 firms. The authors state that they will publish ticker data for additional firms once their data collection is complete. I will remerge the data after the publication to evaluate if my sample contains further noise tickers. If my current sample featured noise ticker, the potential bias would increase the SVI level and work against my predictions.



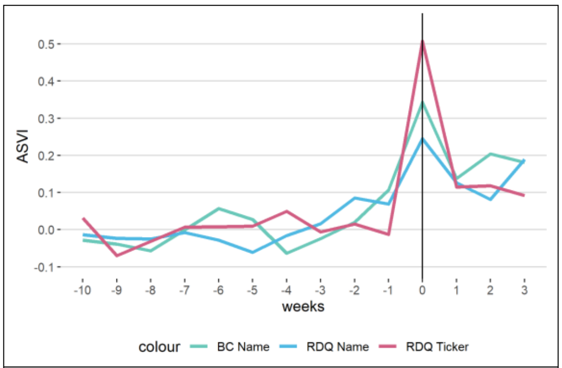
Line	BC Name	BC private
mean w=0	0.322	0.591
t-statistics	5.704***	8.630***

Panel A: Listed and private name comparison



Line	BC Name	BC Ticker
mean w=0	0.322	0.118
t-statistics	5.704***	3.523***

Panel B: Name and Ticker comparison



Line	BC Name	RDQ Name	RDQ Ticker
mean w=0	0.322	0.245	0.507
t-statistics	5.704***	8.854***	13.980***

Panel C: Name and RDQ comparison

Figure 4.2: Cross-sectional average ASVI around Fortune list publication for different search terms
This figure shows the cross-sectional mean Google abnormal search volume index (ASVI) around the week of the list publication based on the weekday of event 1 for the different search terms. Panel A compares the name searches for the public list firms (BC private) with the name searches for private firms and the name searches adding the term "Jobs" (BC private jobs). Panel B compares the name searches for the list firms (BC Name) and the ticker searches (BC ticker) for these firms. Panel C compares the name searches for the list firms (BC Name) with the name (RDQ Name) and ticker (RDQ Ticker) searches for the earnings announcement date (RDQ=report date quarterly). ***, **, * indicate statistical significance at the 1%, 5%, and 10% level, respectively (H_0=0).

Panel A and B of this figure show a significant increase in abnormal search volume for the first event. The ASVI is highest for name searches of private firms (Panel A), while it is lowest for the BC Ticker searches (Panel B) following the publication of the annual list. For the BC list publication, the ASVI based on name searches exceeds the results for the ticker searches. I find that for mainly financial events, i.e., the publication of quarterly earnings, the ASVI based on ticker searches is nearly twice as large as the name search result and also exceeds the results I find for the ticker-based ASVI following the BC list publication (Panel C). The ASVI in the first two panels decreases after w=0 but remains at an increased level in the following three weeks. Figure 4.C4 (Appendix) shows the results of Figure 4.2 for the SVI.

Table 4.4 presents the cross-sectional means and t-tests for the abnormal Google search volume index for both event dates for listed firms and private firms. Additional to Figure 4.2, this table also features the ticker searches excluding the high noise tickers identified by deHaan et al. (2019) and a comparison between event 1 and event 2 on the exact event date. Thus, this table features all supporting steps to alleviate identification concerns. I will discuss the results of each step in turn.[39]

Table 4.4: Supporting identification steps

	(1)	(2)	(3)	(4)
	Name-based	Private Name-based	Ticker-based	Ticker-based (without high noise tickers)
	$H_0 = 0$	$H_0 = 0$	$H_0 = 0$	$H_0 = 0$
Event 1				
N	423	619	423	375
mean	0.322	0.591	0.118	0.114
t-statistics	5.704***	8.630***	3.523***	3.039***
Event 2				
N	411	619	411	364
mean	0.218	0.258	0.115	0.111
t-statistics	4.867***	5.846***	3.170***	2.743***
RDQ comparison				
N	529		529	473
mean	0.245		0.507	0.558
t-statistics	8.854***		13.980***	14.833***

Notes: This table shows the ASVI for the supporting identification steps based on different search terms. Column (1) features results based on firm name searches (as in main results), column (2) features results based on firm ticker searches, and column (3) features ticker results without high-noise tickers. Column (4) shows name searches for the private list firms. RDQ comparison shows a comparison for name and ticker searches with a quarterly earnings announcement date. ***, **, * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

My first supporting identification step compares the ASVI of private and listed firms. Comparable to the results for the listed firms, I find a significantly (1% level) positive increase in stakeholder attention following event 1 for the private firms. Additionally, the ASVI remains statistically significant (1% level) for event 2. While the absolute SVI magnitude is lower for private firms, the ASVI for private firms is nearly twice as high

[39] To facilitate the comparison, I truncate ASVI name observations if the ticker data is missing due to Google Trends data restrictions and the N differs from the other ASVI analysis.

compared to listed firms. The ASVI for listed and private firms become more alike for the second event, yet the private firm ASVI still exceeds the value of the listed firms (Table 4.4, Column (2)).

For my second identification step, I compare ticker- and name-based Google searches for the two event dates. Based on the market reaction results presented in this section, I expect statistically significant ASVI ticker results only for the first event. The ticker ASVI magnitude is less pronounced than for the name searches of listed or private firms (0.118 vs. 0.322 and 0.591 for event 1 and 0.115 vs. 0.218 and 0.258 for event 2). However, I find significantly positive results for the investor attention proxy following both events. Next, I exclude high noise tickers identified by deHaan et al. (2019) from the ticker sample. Column (4) shows the results of this analysis. The ASVI results decrease slightly, but the significance levels remain mainly unchanged compared to column (3). Therefore, the potential noise of the tickers does not seem to affect my results. Consequently, part of the name searches presented in column (1) might be attributable to investors.

In the last step, I compare the name and ticker ASVI from my BC list events with the same proxies following quarterly earnings announcements. In line with the evidence presented in Figure 4.2, I find that the ASVI based on ticker searches largely exceeds the ASVI for the name searches for a mainly financial event. The magnitude for the SVI is comparable for both search terms, while the ASVI is nearly twice as high for ticker searches and even higher once I exclude the potentially noisy tickers. Therefore, it seems that investors use firm tickers rather than firm names when they search for financial information. Based on the results of my identification steps, I am confident that the name-based ASVI is a suitable proxy for stakeholder attention.

4.3.5 Cross-sectional analysis

4.3.5.1 Design

In the next step, I want to provide insights into the firm characteristics that shape stakeholder attention. For all model specifications, I use dependent variables measured on event 1. I employ the following regression model:

(1) $ASVI = \beta_0 + \beta_1 press\ var + \beta_2 Controls + \beta_3 Year\ FE + \beta_4 Firm\ FE + \varepsilon$

The variables of interest for the cross-sectional analyses are the proxies for the public familiarity and the public sentiment towards the list firms, respectively. I use two different variables based on Ravenpack to proxy these firm characteristics.[40] The first version, which captures public familiarity, is based on the amount of firm press coverage in the year prior to the list publication (Twedt (2016)). I measure prior press coverage as the natural logarithm of one plus the number of days with press coverage (*prior_cov*).

For the second version, which captures the sentiment of the public image, I include the sentiment of the news articles in the analysis. *Neg_rel* captures the ratio of negative press articles in the year before the list publication divided by all articles relating to a firm in that year. I identify negative articles based on the RavenPack sentiment score.[41] To capture the additional effect of mostly negative press coverage on the dependent variable, I interact *neg_rel* with *prior_cov*.

In my cross-sectional tests, I use a set of control variables, which are likely associated with media coverage and stakeholder attention. I measure firm attributes at the quarter-end before the list publication. I use the natural logarithm of the market value of equity (*mve*) as a proxy for firm size (Hahn and Kühnen (2013); Fang and Peress (2009)). Following Cahan et al. (2015) and Flammer (2013), I use the natural logarithm of the number of years since the initial firm coverage in Compustat to control for firm age (*age*). To proxy for the growth potential of the list firms, I include the book to market ratio (*btm*). I measure the book to market ratio as the book value of equity at the quarter-end scaled by the market value of equity (Hillert et al. (2014)). To control for information available to investors and stakeholders through other sources than the firm and the press, I include the analyst following of a firm (*AF*). I measure analyst following as the natural logarithm of one plus the mean analyst following of a firm in the year before the list publication. I use the return on assets, measured as net income before extraordinary items scaled by total assets (*roa*) as a proxy for firm profitability (Hawn et al. (2018); Dimson et al. (2015)). Based on my previous results, I include an indicator variable set to one for a top 10 list rank as a control for the list performance (*top10_rank*).

The variable definitions are summarized in Table 4.C5 (Appendix). I winsorize *AF*, *btm*, *mve*, and *roa* at the 1% and 99% level to mitigate the potential effect of outliers. Table

[40] RavenPack features four different news types: (hot) news flashes, full articles, press releases, and tabular material. For my analysis, I only keep full articles, and I eliminate all articles with a relevance score below 75.

[41] For each article, RavenPack provides an event sentiment score between 0 and 100 based on a categorization by financial experts. A score below 50 indicates a negative event, while Ravenpack classifies a score above 50 as positive.

4.5 (Panel A) reports summary statistics for the variables used in this analysis. Table 4.C6 (Appendix) shows the correlation matrix for the different variables. I estimate model (1) using firm and year fixed effects. Additionally, I cluster the standard errors at the firm level.

Table 4.5: Descriptive statistics

Panel A: Cross-sectional analysis list firms						
Variable	mean	p50	sd	min	max	N
asvi	0.34	0.13	1.19	-1.00	13.24	446
prior_cov	4.03	4.25	1.06	0.00	5.38	446
neg_rel	0.28	0.28	0.11	0.00	0.69	446
btm	0.33	0.27	0.25	-0.09	1.49	446
mve	9.48	9.59	1.47	5.10	12.59	446
AF	2.72	2.79	0.59	0.00	3.64	446
age	3.06	3.04	0.53	1.61	4.20	446
roa	0.02	0.02	0.02	-0.04	0.10	446
top10_rank	0.08	0.00	0.28	0.00	1.00	446
Panel B: List firms						
Variable	mean	p50	sd	min	max	N
LnEmp_fines	0.77	0.00	2.96	0.00	16.21	180
btm	0.38	0.28	0.31	-0.06	1.65	180
mve	8.90	8.75	1.42	4.53	11.34	180
AF	2.01	2.46	1.14	0.00	3.16	180
age	2.85	2.77	0.56	1.79	4.22	180
roa	0.02	0.02	0.02	-0.06	0.10	180
Panel C: Control firms						
Variable	mean	p50	sd	min	max	N
LnEmp_fines	1.53	0.00	4.15	0.00	16.64	197
btm	0.45	0.38	0.46	-0.94	2.73	197
mve	8.70	8.45	1.44	5.43	11.34	197
AF	2.07	2.37	0.96	0.00	3.16	197
age	2.79	2.71	0.63	1.10	4.22	197
roa	0.02	0.02	0.02	-0.07	0.09	197

Notes: This table provides descriptive statistics. Panel A shows the descriptive statistics for the cross-sectional analysis. Panel B shows the descriptive statistics for the list firms of the DiD estimation. Panel C shows the descriptive statistics for the control firms of the DiD estimation (N DiD=337 (Panel B and Panel C)). Variables comprise ASVI, prior_cov, neg_rel (press sentiment), btm (Book-to-market), mve (market value of quity), AF (number of analyst forecasts), age, roa (return on assets), and top10_rank_dummy. All variables are defined in Table 4.C5 (Appendix).

4.3.5.2 Results

Table 4.6 shows the OLS regression results based on the press variables that capture the public familiarity and image of a firm. Column (1) shows the results using the number of days with press coverage (*prior_cov*) as the variable of interest. The results show a negative association of the ASVI with previous press coverage (5% level). One explanation for this association might be that stakeholders do not require additional information about firms that they are familiar with and instead search for firms that have been unknown to them before the list publication.

Table 4.6: Cross-sectional analysis of shareholder attention

Variable	(1) ASVI	(2) ASVI	(3) ASVI
prior_cov	-0.889**	-0.831**	-1.460***
	(-2.05)	(-2.07)	(-9.64)
neg_rel		-0.965	-17.009***
		(-0.95)	(-5.29)
prior_cov x neg_rel			3.897***
			(5.66)
mve	0.003	-0.008	-0.058
	(0.02)	(-0.04)	(-0.27)
btm	-1.180***	-1.125**	-0.782*
	(-2.63)	(-2.59)	(-1.89)
AF	-0.349	-0.310	0.059
	(-1.05)	(-0.98)	(0.21)
age	-0.241	-0.290	-0.068
	(-0.34)	(-0.44)	(-0.15)
roa	-2.998	-3.200	-6.238*
	(-1.00)	(-1.05)	(-1.90)
top10_rank	0.620	0.642	0.571
	(1.50)	(1.50)	(1.48)
R^2	0.531	0.532	0.589
N	446	446	446
Cluster	Firm	Firm	Firm
FEs	Firm, Year	Firm,Year	Firm, Year

Notes: Results from estimating an OLS regression using the abnormal search volume index (ASVI) as the dependent variable. Variables comprise ASVI, prior_cov, neg_rel (press sentiment), btm (Book-to-market), mve (market value of quity), AF (number of analyst forecasts), age, roa (return on assets), and top10_rank_dummy. All variables are defined in Table 4.C5 (Appendix). ***, **, * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

In column (2), I add the indicator dummy for mainly negative press coverage (*neg_rel*) and interact it with the amount of prior press coverage (*prior_cov*). In line with column (1), *prior_cov* is significantly negative (1% level) in column (3). Further, the dummy variable for negative press coverage is highly significant (1% level) and negative. The interaction term (*prior_cov x neg_rel*) is positive and significant at the 1% level. As the combined coefficient for *prior_cov* and the interaction term is positive, this means that the increase in stakeholder attention is more pronounced when the firm attracts more negative media attention before the BC list publication. Figure 4.C7 (Appendix) shows a margins plot based on the *neg_rel* distribution and underlines the results of Table 4.6 (column (3)) graphically. Negative press coverage might shape how stakeholders feel towards a firm. The positive information conveyed in the BC list alters the information set of stakeholders regarding this list firm and might, therefore, increase their interest in a firm as a potential employer, as the new information does not match the previous image.

4.4 Labor market reaction and investment effects

In this section, I present and discuss the key aspects of the research design and results for the two different long-term reaction analyses: the stakeholder reactions (labor market reaction) and the firm reactions (investment effects).

4.4.1 Research design

4.4.1.1 Sample

Being featured on the BC list might have two different long-term effect channels for the list firms. First, firms might see an increase in job applications after their first list ranking, as potential employees might be more willing to work for a firm with good employee treatment. Second, the list inclusion might impact the long-term CSR investment decisions of a firm in anticipation of the stakeholder attention. As my sample period starts a few years after the first publication of the BC list in Fortune in 1998, some firms are featured on the list during the entire sample period. For these firms, I am unable to observe labor market reactions or investment effects, as those would have, at least partially, materialized before the start of my sample period. Therefore, I exploit the fact that the first-time list ranking of some firms is dispersed throughout my sample period for both long-term analyses. Examining the first list ranking facilitates the identification of potential labor market reactions and CSR investment effects based on pre/post comparisons at the firm level. The effects should be most pronounced for this subset of observations. There are 45 listed first-time list firms in my sample between 2005 and 2017. These 45 firms are the starting point for both long-term effect samples. I drop several observations from these samples due to data restrictions as outlined in the respective analysis.

4.4.1.2 Labor market reaction

To examine the long-term stakeholder implications of the increase in their attention, I focus on labor market reactions and test whether the number of applications a firm receives increases after being featured on the list for the first time. As stakeholders comprise various groups, the BC list might also affect other stakeholders such as customers and result in product market reactions if they alter their buying behavior. However, the firms on the BC list are from various industries. While some firms have a "business-to-customer" model and might be more exposed to customer reactions (e.g., Starbucks), the list also features business service firms or firms from the oil industry. For these firms, it is less likely to observe product market reactions. Therefore, I focus on the effect that the BC list has on potential employees because the link between the information content of the list and potential stakeholder reactions is strongest for them. Additionally, the information is valuable for employees regardless of the industry or other specific characteristics of the list firms.

My previous analyses show an increase in stakeholder attention following the publication of the BC list. To examine labor market reactions, this increase needs to be based on employee attention instead of other stakeholders. To alleviate concerns that the increase emanates from other stakeholders and to isolate the employees' attention, I collect Google Trends data based on the name of the list firms and add the term "jobs" to the search request (e.g., "Autodesk jobs"). As this search term is rather specific, any increase in Google searches after the list publication should be related to current or future employees.

To test my prediction regarding the application behavior of stakeholders once the firms are featured on the list, I compare the number of job applications in the first list year to the second and third year. For their BC list application, the management of the firms must provide internal items, including the number of applications they received and the total number of employees. Fortune displays these items on its web page accompanying the BC list. I collect the number of applications alongside the number of US, non-US, and total employees from the Fortune web page.[42] If a firm is featured on the list for the first time, the number of applications stated in the Fortune survey refers to the pre-period, i.e., to the time before the firm signaled its good CSR performance to potential employees. In

[42] The web page information is available from 2006 on. Therefore, I exclude 2005 from the sample period for this analysis. As the application data is only available for list firms, it is not possible to estimate a DiD model comparable to the CSR investment effects.

contrast, the number of applications in the second year reflects the level of applications after the first positive signal when stakeholders start updating their beliefs. As it might take some time for the positive signal to materialize in job applications, I also compare the first year to the third year of the list membership.

The data items collected from Fortune are available for private and listed firms. Therefore, I also examine whether the labor market reaction differs for these firms. The variable of interest for the t-tests is *appl_scaled*, for which I divide the number of applications stated on the Fortune web page by the mean number of total employees in both comparison years. For this analysis, I use the 45 first-time firms of the sample. I drop 13 (22) firms because of missing application or scaling data in one of the comparison years. I end up with 32 (23) firms for the second (third) year comparison. In addition to the listed firms, I also examine the labor market reaction for private firms. There are 89 first-time private firms on the BC between 2005 and 2017. I drop 33 (45) firms because of missing data and end up with 56 (44) private firms for the application comparisons.

4.4.1.3 Long-term firm investment effects

As outlined in Section 4.2, I expect an increase in CSR activities once a firm is featured on the list. These long-term investments represent pre-emptive firm reactions to avoid adverse stakeholder reactions after the facilitated monitoring through the BC list. Prior literature such as Lys et al. (2015) and Fiechter et al. (2019) uses CSR scores based on the Asset4 database as a proxy for CSR expenditures and the consequent CSR activities. These scores capture and aggregate various dimensions of firms' CSR activities. For example, the Asset4 social score includes multiple employee aspects such as turnover rates or salary gaps but also features information about product responsibility, like revenues from alcohol or tobacco.

However, I am only interested in a specific aspect of CSR, namely the employee-related CSR activities. Additionally, the scores are also sensitive to the overall coverage of the database. For example, Asset4 largely extended its coverage in 2008, which affects the score of the previously covered firms. As my sample includes years before and after 2008, using Asset4 scores might bias my results. Therefore, the scores are less suited for my analysis. While the database offers various scores, the data available for items that are not part of a score is limited. For example, no item directly captures the extent of CSR investments. Thus, I abstain from using Asset4 data and use the fines a firm must pay due to non-financial corporate misconduct as a proxy for the prior CSR performance instead. To focus on employee-related CSR activities, I only include fines related to employee aspects such as workplace safety. In contrast to CSR scores, the non-financial misconduct fines are less susceptible to changes in firm disclosure about their CSR activities.

I follow Raghunandan (2019) and Heese and Perez-Cavazos (2020) and obtain information about previous non-financial misconduct of BC firms from a novel dataset called Violation Tracker[43] by Good Jobs First. The database features fines for corporate wrongdoing between 2000 and 2018 from 47 federal agencies like the Department of Justice or the Department of Labor Wage and Hour Division. It classifies the fines into different primary-offense categories such as consumer protection, environmental, and bribery. It features the agency that charged the firm and the penalty amount. The database links every fine to the respective agency documents and offers more background information. I collect the employee-related fine record of each sample firm from the database.[44] Appendix B provides more detailed information about the Tracker database and the fines used in my analyses.

To test my prediction empirically, I employ a difference-in-difference model (DiD) to examine the average treatment effect of the list inclusion. *LnEmp_fines*, which is the dependent variable in my DiD model, is the natural logarithm of the total amount (in USD) of fines that a firm had to pay due to employee-related non-financial corporate misconduct in the previous year based on the Violation Tracker data. For the DiD analysis, I estimate the following OLS-regression using standard errors clustered at the firm level:

$$(2)\ LnEmp\_fines = \beta_0 + \beta_1 list\_firms * post + \beta_2 Controls + \beta_3 Year\ FE + \beta_4 Firm\ FE + \varepsilon$$

Using the first list inclusion as treatment provides a clear dichotomy between the list and control firms. To be considered as a treatment list firm, the firm must switch from a non-list firm to a list firm within the sample period. Additionally, I exploit the fact that the first list inclusion is dispersed throughout my sample period. As I expect some time-lag for the fines (e.g., if a court sets the fine), I set the indicator variable *post* to 1 for the year

[43] https://www.goodjobsfirst.org/violation-tracker

[44] The employee related fines in the sample feature the following categories: Family and Medical Leave Act, Employment discrimination, Employment screening violation, Labor relations violation, Wage and hour violation, Workplace safety or health violation.

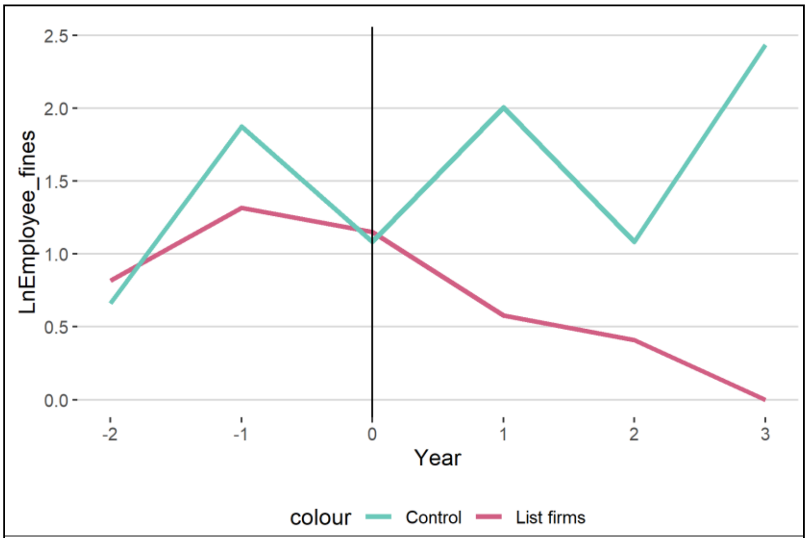
after the first list inclusion.[45] The analysis features three pre and three post years. For each list firm, I require a minimum of five out of six possible observations.[46] Due to this restriction, I drop seven list firms.

Firms in my sample decide to apply for the BC voluntarily. Therefore, they self-select into the treatment (list firms) and the control group. This endogenous treatment decision might lead to a potential self-selection bias. Prior research, such as Bushee et al. (2003) or Butler et al. (2007), addresses this bias using selection models, for example, maximum likelihood estimations or a two-step approach based on Heckman (1979). Both approaches model the firms' decisions to choose a treatment, e.g., the disclosure of quarterly reports and, therefore, control for bias based on a linear combination of both observable and unobservable factors (Tucker (2010)) within one approach.

Though the decision of the firms to apply for the list is endogenous, their final list rank is exogenous, as it is based on the list criteria, which firms cannot influence. Consequently, I need to model the decision to apply or not to apply to the BC list to address the selection bias based on unobservables, regardless of the firm making the list or not. However, Fortune only publishes the 100 best companies, while the name and theoretical rank of the other (approximately) 300 unsuccessful applicants remain private. Due to this data restriction, I abstain from estimating a selection model.

Instead, my research design addresses the potential self-selection bias in two ways. First, I employ a propensity score matching (PSM). The PSM controls for selection bias based on observable factors. The matching procedure creates a sample of comparable list and non-list firms, thus reducing bias from a potential functional form misspecification (Tucker (2010); Shipman et al. (2017)). Additionally, my DiD design addresses potential bias caused by time-invariant unobservables (Lennox et al. (2012)). Yet, my research design would not provide an effective control for the self-selection bias if the selection decision was associated with time-varying unobservables (Lennox et al. (2012)). This restriction limits my ability to draw causal inferences from the DiD analysis.

[45] The results remain mainly unchanged when setting post to 1 in the year of the first list inclusion.
[46] I require a minimum of four observations for the list firms whose first-time list inclusion is at the start (2005) or end (2017) of my sample period.



Year		-2	-1	0	1	2	3	Total
List_firms	without penalty	28	29	30	28	25	22	
	with penalty	3	5	4	4	2	0	
	Total	*31*	*34*	*34*	*32*	*27*	*22*	*180*
Control	without penalty	29	28	31	27	28	26	
	with penalty	2	6	3	7	4	6	
	Total	*31*	*34*	*34*	*34*	*32*	*32*	*197*
Total		62	68	68	66	59	54	377

Figure 4.3: Parallel trends for list and control firms of the DiD estimation
This figure shows the cross-sectional mean of LnEmp_fines for the list and control firms for the DiD sample years, indicating parallel trends before the first-time list membership (treatment). The table at the bottom of the figure documents the sample composition.

For the matched sample, I employ a PSM based on all Compustat firms that have never been on the BC list up to 2018 and the list firms.[47] To obtain the propensity scores, I estimate a probit model of the list firm indicator variable and the covariates I use in the DiD model (Model (2)) based on the fundamentals two years before the list firms are first featured on the BC list.[48] These covariates are the same variables I use as control variables in the cross-sectional analysis except for the list rank (AF, age, roa, mve, and btm). Using a caliper of 0.01, I match the list and control firms without replacement (Leung and Veenman (2018); Shipman et al. (2017)). Due to the caliper, I am unable to match four firms and drop them from the sample. The final sample consists of 34 list and 34 control firms. Table 4.C8 (Appendix) shows that these firms are similar after the matching in the pre-list period because there are no significant differences between the covariates. Table 4.1 (Panel B) features information about the composition of the matched sample. Figure

[47] Excluding firms that are on the list at a later point of time reduces the risk of using a list applicant as a control firm. If I used an applicant as a control firm, this would reduce any effect I might find using the DiD analysis, as these firms might also increase their CSR activities and the difference between treatment and control group would be smaller.
[48] The results remain mainly unchanged when matching based on fundamentals one year before the first list ranking.

4.3 shows the cross-sectional mean of the dependent variable (*LnEmp_fines*) for the list and control firms for the six years of the DiD analysis. The mean increases for both groups two years before the list inclusion and decreases in the year before the treatment. Overall, the figure supports the pre-years parallel trend assumption necessary for the estimation of a DiD model. Table 4.5 (Panel B and C) provide information about the summary statistics for the list and control firms.

4.4.2 Results

4.4.2.1 Labor market reaction

In a first step, I isolate the increase in employee attention by examining an employee-specific search term. Thus, I add "jobs" to the name search term. Table 4.7, column (1) shows the result of this analysis for the listed firms. Comparable to the name searches, I find a significant (1% level) increase in Google searches for both events.[49] The magnitude of the ASVI is similar to the ASVI name magnitude and exceeds the values I find for the ticker searches. Column (2) shows the results for the private firms. Again, I find a significant increase (1% and 5% level, respectively) in Google searches following both events. In line with the name searches, the private mean ASVI for the jobs search exceeds the mean of listed firms. Figure 4.C9 shows the SVI and ASVI comparisons for the search terms outlined above.

Table 4.7: Jobs searches

	(1) Listed firm (Jobs) H_0=0	(2) Private firms (Jobs) H_0=0
Event 1		
N	249	229
mean	0.334	0.500
t-statistics	4.463***	4.620***
Event 2		
N	240	221
mean	0.190	0.581
t-statistics	2.959***	2.200**

Notes: This table shows the ASVI for name searches adding the term "jobs". Column (1) features results for listed firms, while column (2) shows private firms. ***, **, * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

[49] Due to Google Trends data restrictions the N in this analysis is lower compared to the previous one.

In a second step, I examine the number of job applications before and after the first list ranking. Table 4.8 shows the results of the mean comparisons of job applications for the listed firms (Panel A). I find an increase in the number of applications scaled by mean employees. This increase is more pronounced when comparing the first list year with the third one. Nevertheless, neither increase is statistically significant. Panel B shows the results for private firms. Again, there is an increase in applications for both periods. This increase is larger than for listed firms. Yet, the comparisons yield no significant results. Overall, it seems like there is no systematic increase in applications for the firms after being featured on the BC list for the first time. Though the previous analyses indicate an increase in stakeholder attention, the application tests show a lack of stakeholder reaction following their attention. If the information obtained from the BC list does not prompt the stakeholders to act, their Google searches might be based on curiosity rather than information gathering to support their decisions.

Table 4.8: Application data

Panel A: Listed firms					
First and second-year comparison			First and third-year comparison		
	N	Mean		N	Mean
First-year	32	8.935	First-year	23	7.272
Second-year	32	9.826	Third-year	23	10.217
Difference		-0.891	Difference		-2.945
t-statistics		-0.264	t-statistics		-0.959
Panel B: Private firms					
First and second-year comparison			First and third-year comparison		
	N	Mean		N	Mean
First-year	56	7.172	First-year	44	6.588
Second-year	56	8.574	Third-year	44	7.818
Difference		-1.402	Difference		-1.230
t-statistics		-0.971	t-statistics		-0.788

Notes: This table shows the mean comparisons of appl_scaled in the first and second (third) list year for listed (Panel A) and private firms (Panel B). Appl_scaled is the number of applications the firm received scaled by the mean employees in the test years. Both items are based on Fortune list data. ***, **, * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

4.4.2.2 Long-term firm investment effects

Table 4.9 reports the results of the DiD model using an OLS regression. Column (1) estimates a plain model, while columns (2) and (3) add the controls and fixed effects. The main effect of the list firm indicator variable suggests that list firms do not pay less non-financial corporate misconduct fines in the pre-period, which is in accordance with the means presented in Figure 4.3. The estimated average treatment effect of the BC list firms (*list_firm*post*) is significant (10% level and 5% level, respectively) and negative in all model specifications of Table 4.9.

Overall, the results suggest that list firms increase their employee-related CSR activities in anticipation of the increased stakeholder attention and thus must pay fewer non-financial corporate misconduct fines in the following years. This supports my prediction that firms invest in their employee-related CSR performance before making the list for the first time. Additionally, this result suggests that firms are aware of the increased stakeholder attention and the additional accompanying monitoring following the list publication. Thus, they remain at a higher CSR level after making the list to keep sending positive signals to (potential) employees and avoid adverse stakeholder reactions.

Table 4.9: Difference in difference estimation

Variable	(1) LnEmp_fines	(2) LnEmp_fines	(3) LnEmp_fines
list_firm	-0.148	-0.205	
	(-0.22)	(-0.31)	
post	0.569	0.510	
	(1.44)	(1.25)	
list_firm x post	-1.152*	-1.157**	-1.122*
	(-1.96)	(-2.01)	(-1.81)
btm		-0.221	-0.361
		(-0.26)	(-0.38)
mve		0.266	0.830
		(1.01)	(1.33)
AF		-0.106	-0.271
		(-0.55)	(-0.67)
age		-0.006	-4.467
		(-0.01)	(-1.50)
roa		-7.126	-9.059
		(-0.90)	(-0.74)
R^2	0.020	0.033	0.600
N	377	377	377
Cluster	Firm	Firm	Firm
FEs	No	No	Firm, Year

Notes: This table shows the results from estimating a DiD model using the natural logarithm of employee-related fines as the dependent variable based on a matched sample. Variables comprise LnEmp_fines, btm (Book-to-market), mve (market value of quity), AF (number of analyst forecasts), age, and roa (return on assets). All variables are defined in Table 4.C5 (Appendix). ***, **, * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

4.4.3 Robustness test

For the DiD model outlined in this section, I allow firms to have missing years, which leads to a varying number of list and control observations in Table 4.9. To limit concerns that this varying number of observations biases the estimation, I re-estimate equation (2) using a fully balanced sample. Therefore, I drop several list and control firms, which leads to a rather small number of observations (matched sample N=204). Table 4.C10 (Appendix) features the results of this robustness test. The average treatment effect remains negative and slightly significant for all models. Overall, the results of Table 4.C10 support the main results that firms pay fewer non-financial corporate misconduct fines after making the Fortune list for the first time.

Table 4.C11 (Appendix) shows the application t-tests based on the balanced sample. I find an increase in applications when comparing the first and the second list year, though the difference is not statistically significant. Additionally, I examine the first year and the third year of list inclusion. For this comparison, I find an increase in applications, which is slightly statistically significant (10% level). Overall, these results support the main results.

4.5 Conclusion

In this paper, I investigate the link between stakeholder reactions and the publication of non-financial disclosure. I use a sample of Fortune's "100 best companies to work for" firms between 2005 and 2017 to provide evidence on the short-term stakeholder attention following positive, externally validated CSR disclosure. Additionally, I provide evidence on the long-term stakeholder reactions and how firms respond to these actions.

I measure the stakeholder attention using the abnormal Google search volume index and find a highly significant increase in stakeholder attention following the BC list publication. To alleviate concerns about the construct validity of the ASVI, I offer various identification steps based on the timing of the list, the listing status of firms, and different Google search terms. The results of each step support the ASVI as a proxy for stakeholder attention. My results further suggest that stakeholders and investors process non-financial information differently.

In additional analyses, I examine the potential long-term effects of the increased stakeholder attention, namely the labor market reaction and investment effects. I find that there is an increase in job applications to list firms, which is not statistically significant neither for listed nor private firms. Estimating a DiD model, I find that list firms increase their CSR activities once they are featured on the list. Overall, it seems like stakeholders pay attention to the publication of stakeholder related non-financial information but fail to act upon it, whereas the firms invest in their CSR activities in anticipation of the stakeholder attention.

My results are subject to several limitations. First, the results are based on a rather small sample, which limits my possibilities to perform subsample analyses. Yet, as the list firms vary in their characteristics (e.g., private vs. listed firm), the effect on the stakeholder attention might still be generalizable to other settings. Second, my results rely on the construct validity of the ASVI to capture stakeholder attention. Although I try to offer several identification steps to alleviate concerns about the proxy, the ASVI might nevertheless capture some retail investor attention. Third, the analysis of the applications is only based on list firms, as I am unable to observe the application numbers for firms that are not on the list. Therefore, I lack a control sample for this analysis.

Finally, the BC list firms choose to apply for the BC list, which means that the treatment firms self-select themselves into list and control firms. This potential endogeneity arises from the decision of a firm to apply for the list, which I cannot observe. Though my research design addresses this potential self-selection bias, it would not provide an effective control for the self-selection bias if the selection decision was associated with time-varying unobservables (Lennox et al. (2012)). Therefore, the nature of my results does not allow causal inferences.

Despite these limitations, my paper contributes to the literature by providing evidence about how non-investor stakeholders perceive CSR disclosure and how they react to this type of externally validated disclosure of CSR activities. Furthermore, my results contribute to the question of how firms react in anticipation of increased stakeholder attention. As outlined above, it is important to understand the relationship between stakeholders and CSR information to draft effective legislation to meet the information demand of non-investor stakeholders.

4.6 Appendix

Appendix 4.A – Abnormal search volume

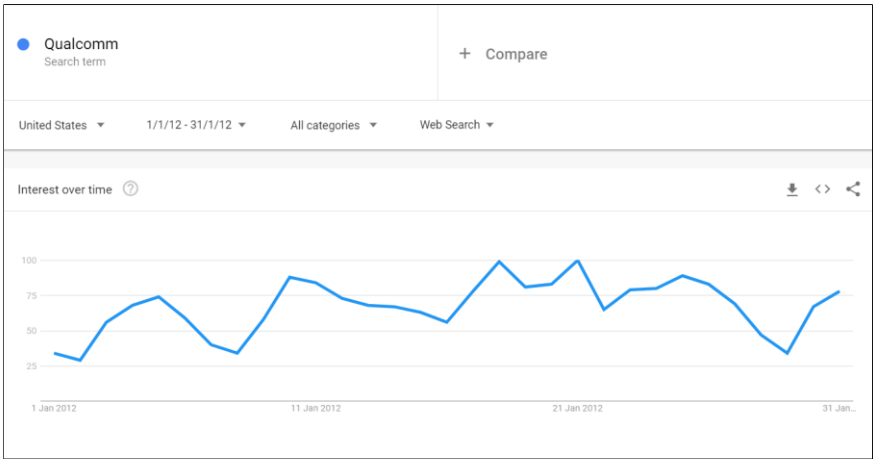
Google Trends provides search term frequency data from January 2004 onwards. Google does not provide the number of searches but a search volume index (SVI). For this index, Google Trends scales the data by the maximum search request volume during the request window. Google provides the SVI data on a daily, weekly, and monthly basis, depending on the length of the request window. In case of very few Google searches, Google truncates the data and provides an SVI of 0.

Prior literature uses firm tickers to capture the investor reaction following financial news (e.g., Drake et al. (2012); Da et al. (2011)). As I want to capture the stakeholder reaction following the BC list publication, I follow Madsen and Niessner (2019) and use Google Trends data based on firm name searches. Following deHaan et al. (2019), I measure ASVI as the SVI for firm i on the day of the list publication less the mean SVI for the firm on the same weekday over 10 weeks prior to the list publication, scaled by the mean SVI for the firm on the same weekday over 10 weeks prior to the list publication.

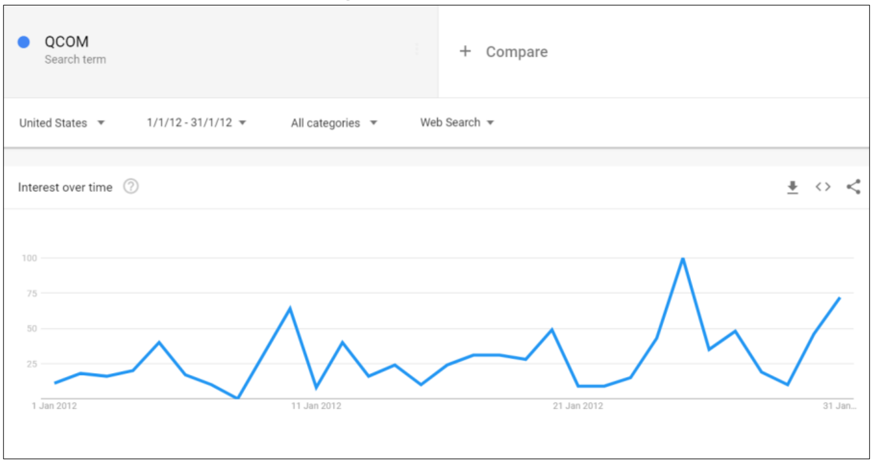
Over the sample period from 2005 to 2017, the data is only available on a monthly level. To calculate the above mentioned ASVI, I need daily Google SVI data for the entire sample period. I follow deHaan et al. (2019) to obtain daily SVI for the sample period. First, I download monthly SVI data for 2004 to 2017. Second, I download daily SVI for each month in which the firm is part of the list, starting three months before the list publication (list window). Finally, I need to rescale the data. I multiply the daily data obtained in step 2 by the monthly data collected in step 1 and rescale it with the maximum monthly SVI overserved for a list firm during the list window. Consequently, each list firm has a maximum value of 100 during its list window.

To collect the SVI data, I prepare the firm names. In a first step, I check for significant names changes and exclude these firms from the sample. If firms only partly changed their names (e.g., from "Deloitte" to "Deloitte & Touche"), I keep the firm in the sample using the original name. In the next step, I remove all special characters from firm names like "&" and remove typical firm name endings like "Group" or "Company". With these cleared names, I collect the monthly and daily SVI. I remove firms from the sample if the SVI for the firm on the same weekday as the list publication over 10 weeks before the list publication is zero more than 7 times. For both events 1 and 2, I use the SVI for the firm on the respective weekday over 10 weeks before event 1 to calculate ASVI because I do not want the stakeholder reaction after event 1 to impact the ASVI for event 2.

Panel A: Qualcomm name search



Panel B: Qualcomm ticker search



Panel C: Qualcomm name and ticker search comparison

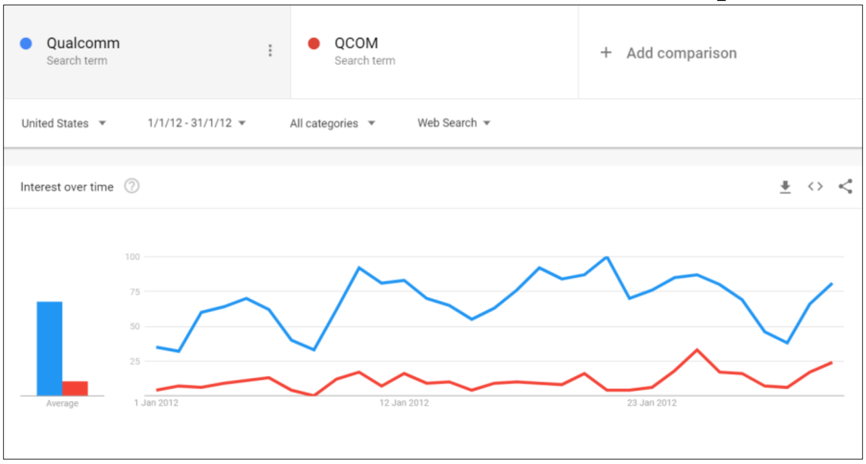


Figure 4.A1: Example of Google Trends data

This figure shows an example for Google name searches (Panel A), Google ticker searches (Panel B) and comparison for both search terms (Panel C).

Appendix 4.B – Violation tracker database

Appendix 4.B1: Corporate misconduct of list firms between 2005 and 2017

Primary Offense	Freq.	Perc.	Cumm.	Mean amount (in $ Mio.)	Over $1 Mio.
Fair Credit Reporting Act violation	1	0.04	0.04	2.300	1
False Claims Act and related	77	3.28	3.33	40.329	68
Family and Medical Leave Act	**11**	**0.47**	**3.8**	**0.021**	**0**
Foreign Corrupt Practices Act	18	0.77	4.56	10.796	15
HHS civil monetary penalties	1	0.04	4.61	0.204	0
Servicemembers Civil Relief Act	1	0.04	4.65	12.000	1
accounting fraud or deficiencies	6	0.26	4.9	75.250	6
anti-money-laundering deficiencies	3	0.13	5.03	58.867	3
aviation consumer protection violation	23	0.98	6.01	10.105	3
aviation safety violation	678	28.91	34.93	0.021	1
banking violation	15	0.64	35.57	7.108	8
bankruptcy professional violation	1	0.04	35.61	2.350	1
benefit plan administrator violation	33	1.41	37.01	15.654	21
campaign finance violation	1	0.04	37.06	0.217	0
civil contempt	1	0.04	37.1	2.000	1
consumer protection violation	68	2.9	40	11.100	25
data submission deficiencies	2	0.09	40.09	0.848	1
discriminatory practices	4	0.17	40.26	0.351	0
drug or medical equipment safety violation	10	0.43	40.68	52.494	10
economic sanction violation	9	0.38	41.07	0.074	0
employment discrimination	**52**	**2.22**	**43.28**	**2.393**	**13**
employment screening violation	**4**	**0.17**	**43.45**	**0.483**	**0**
environmental violation	475	20.26	63.71	2.578	40
excise tax violation	3	0.13	63.84	0.186	0
export control violation	5	0.21	64.05	4.985	1
federal leasing royalty violation	3	0.13	64.18	1.104	1
financial institution supervision failures	4	0.17	64.35	8.590	4
foreign exchange market manipulation	1	0.04	64.39	120.000	1
fraud	1	0.04	64.43	8.250	1
illicit political contributions	1	0.04	64.48	12.000	1
investor protection violation	15	0.64	65.12	239.300	6
kickbacks and bribery	12	0.51	65.63	34.615	9
labor relations violation	**108**	**4.61**	**70.23**	**0.036**	**0**
mortgage abuses	4	0.17	70.41	101.710	4
motor vehicle safety violation	5	0.21	70.62	0.020	0
off-label/unapproved promotion medical prod.	30	1.28	71.9	285.580	27
offshore drilling violation	9	0.38	72.28	0.028	0
pipeline safety violation	1	0.04	72.32	0.037	0
price-fixing or anti-competitive practices	10	0.43	72.75	17.052	5
privacy violation	20	0.85	73.6	14.212	13
product safety violation	9	0.38	73.99	1.109	3
railroad safety violation	55	2.35	76.33	0.009	0
securities issuance or trading violation	10	0.43	76.76	17.917	8
tax violations	13	0.55	77.31	0.034	0
telecommunications violation	29	1.24	78.55	10.721	14
tobacco litigation	1	0.04	78.59	0.100	0
toxic securities abuses	5	0.21	78.81	1797.600	5
wage and hour violation	**198**	**8.44**	**87.25**	**5.417**	**76**
workplace safety or health violation	**299**	**12.75**	**100**	**0.014**	**0**

Notes: This table provides information about the corporate misconduct of list firms between 2005 and 2017 (bold letters indicate employee offenses).

Panel A: Starbucks Penalty record in 2011

Company	Parent	Parent Major Industry	Primary Offense Type	Year	Agency	Penalty Amount
Starbucks	Starbucks	restaurants and foodservice	wage and hour violation	2011	private lawsuit-federal	$1,564,000
Starbucks Coffee Co.	Starbucks	restaurants and foodservice	employment discrimination	2011	EEOC	$75,000
Starbucks Coffee Company	Starbucks	restaurants and foodservice	labor relations violation	2011	NLRB	$5,500

Panel B: Individual record Starbucks

Violation Tracker Individual Record

Company: Starbucks Coffee Co.

Current Parent Company: Starbucks

Penalty: $75,000

Year: 2011

Date: August 18, 2011

Primary Offense: employment discrimination

Secondary Offense: Americans with Disabilities Act

Violation Description:
Disabilities

Level of Government: federal

Action Type: agency action

Agency: Equal Employment Opportunity Commission

Civil or Criminal Case: civil

Facility State: Texas

HQ Country of Parent: USA

HQ State of Parent: Washington

Ownership Structure of Parent: publicly traded

Major Industry of Parent: restaurants and foodservice

Specific Industry of Parent: restaurants

Source of Data:
Https://www.eeoc.gov/eeoc/newsroom/release/8-18-11.cfm (archived copy)

Notes:
Description field shows type of discrimination that EEOC had alleged.

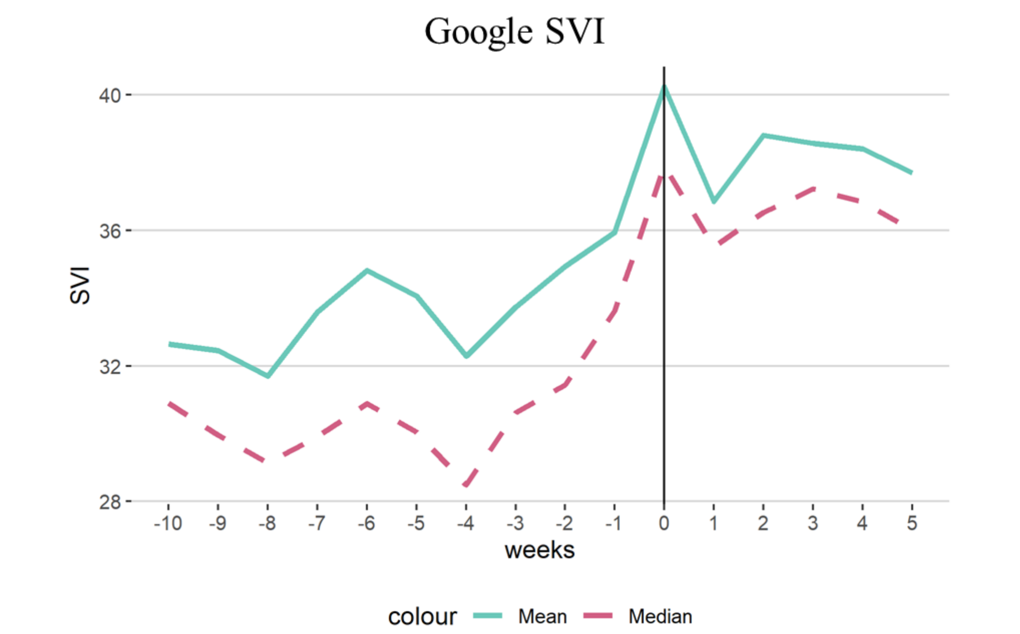
Source Notes: If an online information source is not working, check the Violation Tracker Data Sources page for an updated link.

Parent company note: Parent-subsidiary relationship is current as of the most recent revision listed in the Update Log.

Figure 4.B2: Example of Violation Tracker data
Figure B1 shows an example of the Violation Tracker data. Panel A shows the penalty record for Starbucks, while Panel B shows an individual record from the ones listed in Panel A.

Table B1 and Figure B1 feature information from the Violation Tracker by Good Jobs First. The database features fines for corporate wrongdoing between 2000 and 2018 from 47 federal agencies like the Department of Justice or the Department of Labor Wage and Hour Division. The database classifies the fines into different primary-offense categories like shown in Table B1. It features the agency that charged the firm and the penalty amount (Figure B1, Panel A). The database links every fine to the respective agency documents and offers more background information (Figure B1, Panel B).

Appendix 4.C – Additional tables and figures



$$SVI\ t_0 = 40.24 \quad SVI\ t_{10\ weeks} = 33.62 \quad Diff = 6.62^{***}$$

Figure 4.C1: Cross-sectional average SVI around Fortune list publication
This figure shows the cross-sectional mean and median Google search volume index (SVI) around the week of the list publication based on the weekday of event 1. Week 0 (vertical line) is the week of the list publication based on the first mention of the publication. The sample is based on list publications between 2005 and 2017 (N=446). ***, **, * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 4.C2: Overall Market Reaction event dates (positive list ranking)

Panel A: Market reaction event 1 (date first disclosure)

(1) Event Date	(2) List year	(3) N	(4) Cum.ret. (-1/+1)	(5) S&P500 (-1/+1)	(6) MAR (4)-(5)	(7) t-statistic (H_0= 0)
10jan2005	2005	29	0.000	-0.004	0.004	1.291
09jan2006	2006	23	0.019	0.013	0.006	1.298
08jan2007	2007	26	-0.004	-0.004	0.000	0.029
22jan2008	2008	23	0.033	0.004	0.028	1.980*
22jan2009	2009	20	0.034	0.034	0.000	-0.002
21jan2010	2010	16	-0.043	-0.052	0.009	1.404
20jan2011	2011	18	-0.011	-0.009	-0.002	-0.379
19jan2012	2012	16	0.031	0.017	0.014	1.880*
16jan2013	2013	19	0.012	0.007	0.005	0.762
16jan2014	2014	15	0.002	0.000	0.002	0.386
05mar2015	2015	18	-0.019	-0.017	-0.002	-0.399
03mar2016	2016	16	0.010	0.011	-0.001	-0.111
09mar2017	2017	22	0.003	0.002	0.001	0.933
Mean		***261***	***0.006***	***0.001***	***0.005***	***2.306*****

Panel B: Market reaction event 2 (issue date)

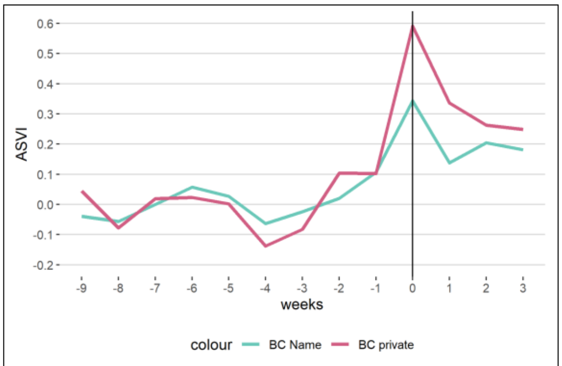
(1) Event Date	(2) List year	(3) N	(4) Cum.ret. (-1/+1)	(5) S&P500 (-1/+1)	(6) MAR (4)-(5)	(7) t-statistic (H_0= 0)
24jan2005	2005	28	-0.018	-0.006	-0.012	-2.113**
23jan2006	2006	22	-0.012	-0.014	0.002	0.463
22jan2007	2007	26	-0.002	0.001	-0.003	-0.557
04feb2008	2008	22	-0.025	-0.030	0.006	1.078
02feb2009	2009	18	-0.012	-0.008	-0.004	-0.239
08feb2010	2010	15	0.016	0.007	0.009	1.404
07feb2011	2011	17	0.022	0.013	0.009	1.638
06feb2012	2012	16	0.022	0.016	0.005	1.253
04feb2013	2013	19	0.011	0.009	0.002	0.617
03feb2014	2014	15	-0.013	-0.022	0.009	1.518
16mar2015	2015	18	0.015	0.004	0.011	1.879*
15mar2016	2016	16	-0.010	0.003	-0.013	-1.068
15mar2017	2017	22	0.006	0.003	0.003	1.102
Mean		***254***	***-0.001***	***-0.002***	***0.001***	***0.555***
t-statistic (event 1 = event 2)					1.279	

Notes: This table provides an analysis of the overall market reaction to the first disclosure of the BC list (Panel A) and the publication of the Fortune issue containing the BC list (Panel B). Column (4) reports the three-day cumulative return of sample firms centered on each event date. Column (5) presents the three-day cumulative return for the S&P500 index. MAR in column (6) is the difference between column (4) and column (5). I test whether the mean return for three-day MAR (column (6)) is significantly different from zero (H_0=0), which is denoted by t-statistic in column (7). Panel B additionally compares the results for event 1 and event 2). All variables are defined in Table 4.C5. ***, **, * indicate statistical significance at the 1%, 5%, and 10% level, respectively. The sample is limited to firms with a ranking improvement.

Table 4.C3: Overall Market Reaction event dates (negative list ranking)

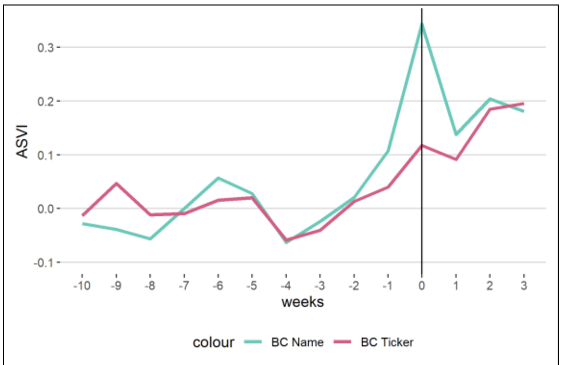
Panel A: Market reaction event 1 (date first disclosure)						
(1) Event Date	(2) List year	(3) N	(4) Cum.ret. (-1/+1)	(5) S&P500 (-1/+1)	(6) MAR (4)-(5)	(7) t-statistic (H_0= 0)
10jan2005	2005	14	-0.014	-0.004	-0.010	-1.789*
09jan2006	2006	17	0.013	0.013	0.000	0.078
08jan2007	2007	14	-0.002	-0.004	0.002	0.513
22jan2008	2008	12	0.024	0.004	0.019	0.781
22jan2009	2009	10	0.038	0.034	0.004	0.333
21jan2010	2010	13	-0.052	-0.052	0.000	-0.010
20jan2011	2011	15	-0.015	-0.009	-0.006	-0.887
19jan2012	2012	14	0.024	0.017	0.007	0.762
16jan2013	2013	12	0.017	0.007	0.010	1.813*
16jan2014	2014	14	0.014	0.000	0.014	1.838*
05mar2015	2015	13	-0.015	-0.017	0.002	0.359
03mar2016	2016	13	0.017	0.011	0.007	1.200
09mar2017	2017	11	0.007	0.002	0.006	0.891
Mean		***172***	***0.004***	***0.000***	***0.004***	***1.591***
Panel B: Market reaction event 2 (issue date)						
(1) Event Date	(2) List year	(3) N	(4) Cum.ret. (-1/+1)	(5) S&P500 (-1/+1)	(6) MAR (4)-(5)	(7) t-statistic (H_0= 0)
24jan2005	2005	13	-0.025	-0.006	-0.019	-1.734
23jan2006	2006	17	-0.013	-0.014	0.001	0.192
22jan2007	2007	12	0.012	0.001	0.011	1.799*
04feb2008	2008	12	0.022	-0.030	0.052	1.182
02feb2009	2009	10	-0.017	-0.008	-0.009	-1.287
08feb2010	2010	13	0.011	0.007	0.003	0.783
07feb2011	2011	15	0.019	0.013	0.005	0.988
06feb2012	2012	13	0.017	0.016	0.001	0.151
04feb2013	2013	12	0.027	0.009	0.018	0.847
03feb2014	2014	14	-0.021	-0.022	0.001	0.142
16mar2015	2015	13	0.009	0.004	0.005	0.920
15mar2016	2016	13	0.001	0.003	-0.001	-0.240
15mar2017	2017	11	0.011	0.003	0.008	2.098*
Mean		***168***	***0.004***	***-0.002***	***0.006***	***1.406***
t-statistic (event 1 = event 2)					-0.246	

Notes: This table provides an analysis of the overall market reaction to the first disclosure of the BC list (Panel A) and the publication of the Fortune issue containing the BC list (Panel B). Column (4) reports the three-day cumulative return of sample firms centered on each event date. Column (5) presents the three-day cumulative return for the S&P500 index. MAR in column (6) is the difference between column (4) and column (5). I test whether the mean return for three-day MAR (column (6)) is significantly different from zero (H_0=0), which is denoted by t-statistic in column (7). Panel B additionally compares the results for event 1 and event 2). All variables are defined in Table 4.C5. ***, **, * indicate statistical significance at the 1%, 5%, and 10% level, respectively. The sample is limited to firms with a ranking deterioration.



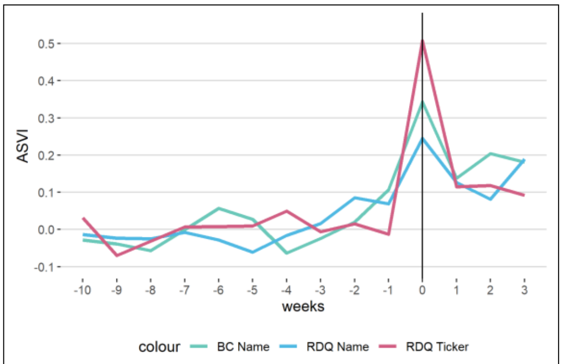
Line	BC Name	BC private
mean w=0	0.322	0.591
t-statistics	5.704***	8.630***

Panel A: Listed and private name comparison



Line	BC Name	BC Ticker
mean w=0	0.322	0.118
t-statistics	5.704***	3.523***

Panel B: Name and Ticker comparison



Line	BC Name	RDQ Name	RDQ Ticker
mean w=0	0.322	0.245	0.507
t-statistics	5.704***	8.854***	13.980***

Panel C: Name and RDQ comparison

Figure 4.C4: Cross-sectional average SVI around Fortune list publication for different search terms

This figure shows the cross-sectional mean Google search volume index (SVI) around the week of the list publication based on the weekday of event 1 for the different search terms. Panel A compares the name searches for the public list firms (BC private) with the name searches for private firms. Panel B compares the name searches for the list firms (BC Name) and the ticker searches (BC ticker) for these firms. Panel C compares the name searches for the list firms (BC Name) with the name (RDQ Name) and ticker (RDQ Ticker) searches for the earnings announcement date (RDQ=report date quarterly). ***, **, * indicate statistical significance at the 1%, 5%, and 10% level, respectively (H_0=0).

Table 4.C5: Variable definitions

Variable	Definition	Data source
Appl_scaled	Number of applications the firm received in the previous year scaled by the mean number of employees	BC list
ASVI	SVI based on name searches on the list publication date minus the ten-week SVI average on the same weekday scaled by the ten-week average (for detailed calculation see Appendix A)	Google Trends
LnEmp_fines	Natural logarithm of employee-related CSR penalties in the year before the list publication	Violation Tracker
MAR	Market adjusted returns [-1/+1], adjusted for the S&P500 market index	CRSP
neg_rel	negative press articles about a firm in the year prior to the list publication scaled by all articles	Ravenpack
post	Indicator variable set to one in the year after the first list inclusion	BC list
prior_cov	Natural logarithm of one plus the number of days with press coverage in the year before list publication	Ravenpack
SVI	Google Trends search volume index	Google Trends
List_firm	Indicator variable set to one if the firm is featured on the BC list (treatment firms) and zero otherwise (control firms)	BC list
Control variables		
age	Natural logarithm of the number of years since Compustat coverage initiation	Compustat
AF *	Natural logarithm of 1 plus the mean analyst following in the year before the list publication	IBES
bad_news	Indicator variable set to one if the firm reported negative Income before extraordinary items in the last quarter	Compustat
btm *	Book value of equity scaled by the market value of equity at the quarter-end [ceqq/(chsoq*prccq)]	Compustat
mve *	Natural logarithm of the market value of equity at the quarter-end [chsoq*prccq]	Compustat
top10_rank	Indicator variable set to one if the firm belongs to the top 10 firms in the list year	BC list
roa *	Income before extraordinary items scaled by total assets	Compustat

Notes: This table shows variable definitions. * denotes the variable is winsorized at the 1% and 99% level due to outliers.

Table 4.C6: Correlation matrix

	ASVI	prior_cov	neg_rel	mve	btm	AF	age	roa	top10_rank
ASVI	1.000								
prior_cov	-0.212***	1.000							
neg_rel	-0.194***	0.613***	1.000						
mve	-0.032	0.065	0.004	1.000					
btm	0.014	0.003	0.030	-0.222***	1.000				
AF	-0.071	0.143**	0.091	0.653***	-0.171***	1.000			
age	-0.057	0.162***	-0.022	0.283***	-0.099*	0.063	1.000		
roa	-0.084	-0.080	-0.068	0.081	-0.265***	-0.023	0.148**	1.000	
top10_rank	0.268***	-0.062	0.074	0.147**	-0.056	0.121*	-0.142**	-0.021	1

Notes: This table provides details on the correlation between the independent variables. All variables are defined in Table 4.C5 (Appendix). ***, **, * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

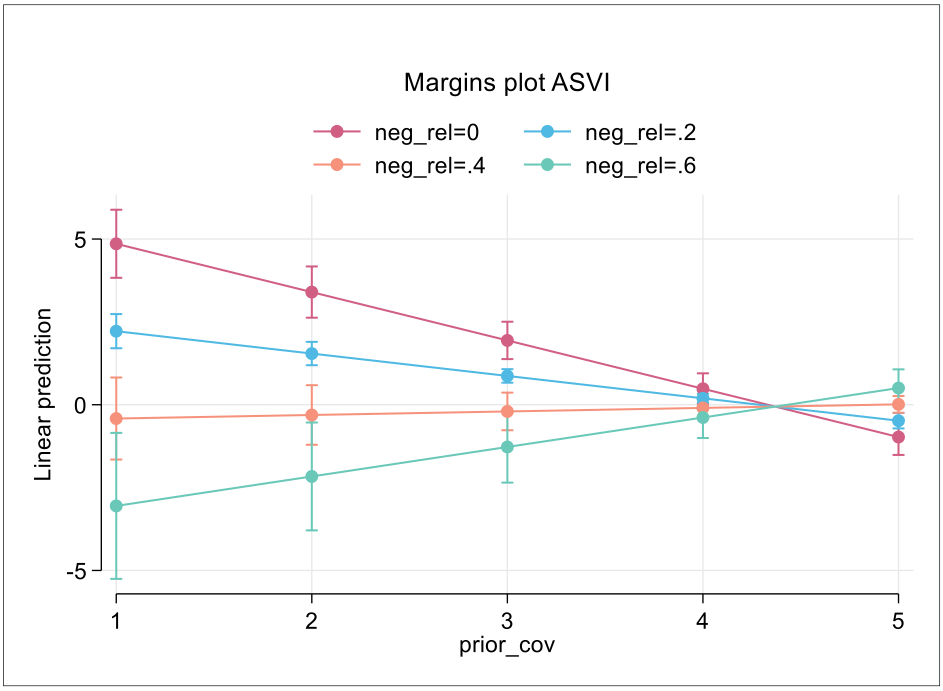


Figure 4.C7: Interaction plot prior press coverage and negative press sentiment for ASVI
This figure shows the margins plots of model (1), including the interaction term of prior_cov x neg_rel ASVI as the dependent variable in an OLS regression.

Table 4.C8: Propensity score matching covariates

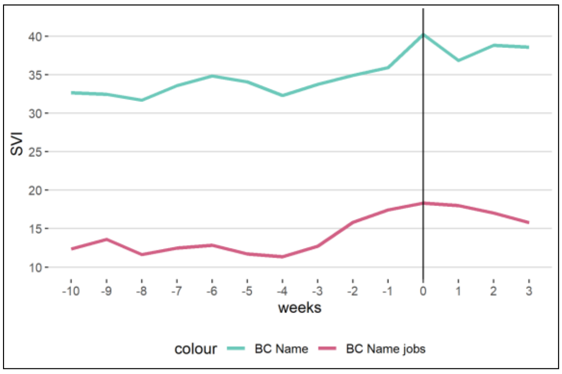
Panel A: Probit model

Variable	(1) Pre-period Indicator variable (1=Treatment, 0=Control)	
LnEmp_fines	-0.015	(-0.45)
btm	-0.445	(-0.81)
mve	0.077	(0.59)
AF	-0.226	(-1.29)
age	0.209	(0.72)
roa	-1.709	(-0.26)
R^2	0.024	
N	198	
Cluster	Firm	
FEs	Year	

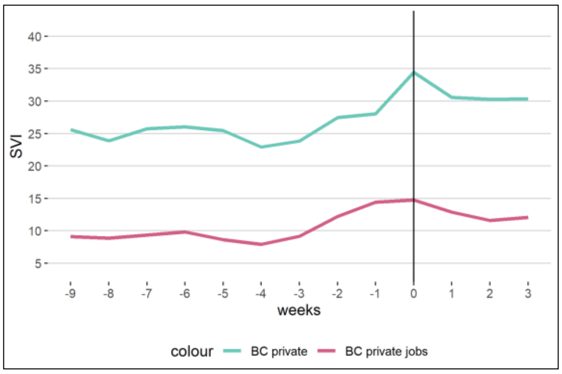
Panel B: Mean differences

	Mean value		Difference (1) vs. (2)	
Variable	(1) Treated	(2) Controls	Difference	t-statistic
LnEmp_fines	1.10	1.22	-0.12	0.23
btm	0.35	0.38	-0.03	0.71
mve	8.84	8.72	0.12	-0.58
AF	2.02	2.21	-0.18	1.36
age	2.79	2.70	0.09	-0.94
roa	0.02	0.02	0.00	-0.12

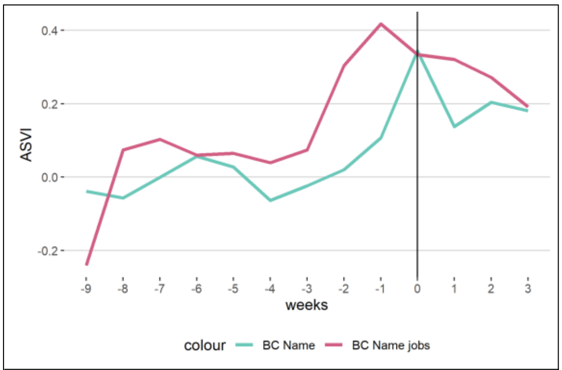
Notes: Results from estimating a probit regression with an indicator variable (1 = Treatment and 0 = Control) as the dependent variable. Variables comprise LnEmp_fines, btm (Book-to-market), mve (market value of equity), AF (number of analyst forecasts), age, and roa (return on assets). All variables are defined in Table 4.C5 (Appendix). Sample details are described in Table 4.1. ***, **, * indicate statistical significance at the 1%, 5%, and 10% level.



Panel A: SVI Name/Jobs comparison

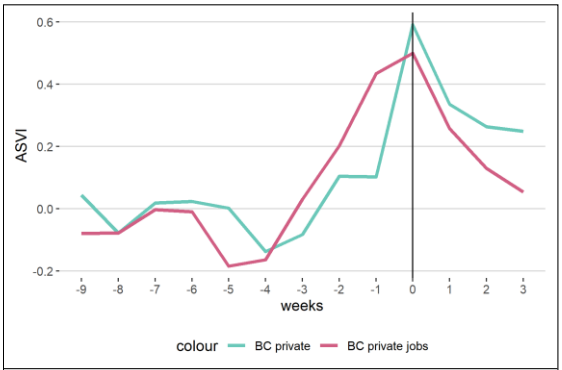


Panel B: SVI Private Name/Jobs comparison



Line	BC Name	BC Name jobs
mean w=0	0.322	0.334
t-statistics	5.704***	4.463***

Panel C: ASVI Name/Jobs comparison



Line	BC private	BC private jobs
mean w=0	0.591	0.500
t-statistics	8.630***	4.607***

Panel D: ASVI Private Name/Jobs comparison

Figure 4.C9: Cross-sectional average SVI for “jobs” searches

This figure shows the cross-sectional mean Google search volume index (SVI) and abnormal search volume index (ASVI) around the week of the list publication based on the weekday of event 1 for the different search terms. Panel A shows the SVI for name searches for listed firms and “jobs” to the search term. Panel B shows the SVI results for name and “jobs” searches for private firms. Panel C shows the ASVI for name searches for listed firms and “jobs” to the search term. Panel D shows the ASVI results for name and “jobs” searches for private firms.

Table 4.C10: Difference in difference estimation (balanced sample)

Variable	(1) LnEmp_fines	(2) LnEmp_fines	(3) LnEmp_fines
list_firms	0.310	0.103	
	(0.35)	(0.11)	
post	0.828	0.402	
	(1.30)	(0.71)	
list_firms x post	-1.709**	-1.573**	-2.078**
	(-2.12)	(-1.98)	(-2.01)
btm		0.761	1.653
		(0.65)	(1.22)
mve		-0.329	0.483
		(-1.07)	(0.61)
AF		0.229	0.440
		(1.17)	(1.17)
age		1.590	-5.105
		(1.16)	(-0.93)
roa		-0.020	8.645
		(-0.00)	(0.94)
R^2	0.026	0.087	0.553
N	204	204	204
Cluster	Firm	Firm	Firm
FEs	No	No	Firm, Year

Notes: This table shows the results from estimating a DiD model using the natural logarithm of employee-related fines as the dependent variable based on a matched sample. Variables comprise LnEmp_fines, btm (Book-to-market), mve (market value of quity), AF (number of analyst forecasts), age, and roa (return on assets). All variables are defined in Table 4.C5 (Appendix). ***, **, * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 4.C11 Application data (balanced sample)

First and second-year comparison			First and third-year comparison		
	N	Mean		N	Mean
First-year	16	6.196	First-year	13	6.100
Second-year	16	9.424	Third-year	13	13.251
Difference		-3.228	Difference		-7.150
t-statistics		-1.530	t-statistics		-2.028*

Notes: This table shows the mean comparisons of appl_scaled in the first and second (third) list year for listed firms. Appl_scaled is the number of applications the firm received scaled by the mean employees in the test years. Both items are based on Fortune list data. ***, **, * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

5 Conclusion

5.1 Summary of major findings and implications of the thesis

Based on three empirical studies, this thesis addresses and extends two general streams of literature: first, the stream on the financial press an information intermediary (e.g., Drake et al. (2014); Bushee et al. (2010); Guest (2018)), and second, the stream on the link between stakeholder theory and non-financial disclosure (e.g., Madsen and Rodgers (2015); Hombach and Sellhorn (2018)).

Contributing to the literature related to the financial press as an information intermediary, the first and the second study differentiate the information dissemination and information production role of the financial press. To provide evidence on these roles, both studies examine different attributes of financial press articles. To that end, the *first study* provides explorative insights into the coverage decision of (financial) news outlets and the content that the articles supply for the readers. Focusing on quarterly earnings announcements of S&P 500 firms, the study performs a content analysis collecting all quantitative items from the articles. The evidence presented in the first study suggests that the press assumes both an information dissemination and production role. While the dissemination focuses on key performance indicators such as earnings or revenues, the information production features background information such as M&A activities. Probit regressions reveal that both the coverage and information production are associated with firm fundamentals such as firm size and do not vary substantially between specialized and general-interest newspapers. Analyses of the quotes in the articles suggest that the information production is not limited to quantitative items but extends to the textual features of the articles.

To glean additional insights into how the financial press produces information based on the textual features of the articles and what kind of information is produced, the *second study* of this thesis provides evidence using a topic modeling approach (LDA). To that end, the study applies the LDA model to press articles covering quarterly earnings and the corresponding firm press releases. The study identifies four measures of information production by comparing the topics in the articles and the earnings announcements. The first measure investigates original article content. On average, the journalists add 1.5 topics per article. Second, the study identifies topics that journalists decide not to report in their articles. On average, the articles omit 5.5 topics. Multivariate analyses show that adding and omitting topics helps investors analyze earnings news. Third, journalists produce additional information by (de-) emphasizing specific topics. At last, journalists produce information based on how they phrase their content. Articles have, on average, a more negative sentiment than the earnings announcements. The sentiment of added topics is significantly more positive compared to joint topics. However, multivariate analyses suggest that only the sentiment of joint topics helps investors analyze earnings news, while there is no significant association for the sentiment of added topics.

Taken together, the first two studies of this thesis support that the financial press creates value for its readers based on information production and exceeds a mere entertainment role. The results imply that information production is a multi-attribute concept because articles produce information on the quantitative, on the topic, and on the sentiment level. Interestingly, both studies, though using a different research approach, show that journalists frequently include quotes (statements) from firm executives, analysts, or other experts to provide additional content. This result suggests that journalists can produce this information due to their personal ties with the quoted persons, which is in line with how journalists describe their work according to Call et al. (2018). Additionally, it seems like what journalists write in their articles is more important than how they say it.

Contributing to the second stream of literature, the *third study* examines if stakeholders pay attention to non-financial information disseminated in the press and how this attention might have long-term consequences for firms based on stakeholder reactions and their own actions. To that end, the study examines the Fortune's "100 best companies to work for" list as a specific form of non-financial disclosure. Using Google searches as a proxy for the stakeholder attention, the study finds an increase in stakeholder attention following the first disclosure of the information and its wider dissemination in the Fortune issue. Additionally, the list provides valuable information content for investors when the information is first disclosed. However, the increase in attention is not followed by corresponding stakeholder actions in the form of job applications. Based on a difference-in-difference model, the study finds an increase in employee-related CSR activities of the firms once they are featured on the list.

Overall, the results imply that stakeholders pay attention to the publication of stakeholder related non-financial information in Fortune magazine. Therefore, the value of information dissemination through the media extends to non-financial information and non-investor stakeholders. Consequently, the press is a suitable channel for firms to disseminate their non-financial disclosure. However, the results also imply that stakeholder awareness is not enough to trigger stakeholder reactions. In contrast, the results support prior literature that firms react in anticipation of increased stakeholder attention and adapt their CSR activities to prevent adverse stakeholder reactions.

5.2 Limitations

The findings of this thesis are subject to several limitations, as outlined in the respective chapters. First, all three studies are subject to endogeneity concerns. For the first and second study, the decision to write an article lies with the news outlets and is therefore endogenous. The characteristics that shape the coverage decision might also change over time. For example, the WSJ merged its news desk with the Dow Jones Newswire in 2013. Consequently, this change in resources might have impacted the coverage decisions of the WSJ (e.g., due to free resources after eliminating overlapping coverage), which might also affect the information production of the outlet. As this restructuring occurred during the sample period of both study 1 and study 2, the merge might affect the results of both studies. For the third study, the endogeneity arises from the self-selection of the firms because they decide to apply to the Fortune list. Endogeneity issues are common in the accounting literature, especially in the literature stream related to the financial press. Overall, these concerns keep the papers from drawing causal inferences. Yet, Glaeser and Guay (2017) point out that research designs that allow the identification of causal relationships are rare and that we should, therefore, use studies on particular research fields using various samples and methodologies as input to update our priors about the theory being tested in a Bayesian manner.

Second, the studies rely on the validity of the empirical constructs. In the second study, most test variables are based on the LDA model and the basic model assumptions of this algorithm, such as the irrelevance of the order of words in a document or the documents within the sample. In the third study, the construct validity of the Abnormal Search Volume Index (ASVI) is crucial. Though the study offers several identification steps, it is still possible that retail investor searches bias the estimates. Additionally, I cannot directly observe the CSR activities of the firms and thus need to rely on the observable non-financial corporate misconduct fines. Third and related, the studies, especially the third one, are subject to data constraints. For example, I am unable to observe the application data for firms that apply to the Fortune list but do not end up among the 100 best firms. Similarly, I cannot observe which firms choose to apply for the list in the first place.

Fourth, the results of the studies might not generalize to other settings. Study 1 and study 2 use a US setting. This might limit the generalizability of the results regarding the financial press in other countries. Additionally, both papers focus on S&P 500 firms. These firms are large and well-known. Thus, their information environment might be different, which might also impact the information production of the press for these firms. The third study examines a specific form of non-financial disclosure. Therefore, the results might not translate to more traditional disclosure types like CSR disclosure in annual reports. Similarly, the long-term results might not transfer to other stakeholder groups, which are not directly affected by the disclosed non-financial information.

Finally, some research design choices are subjective. For example, the content analysis of the first study relies on the classification of the researcher. Additionally, the topic names in study 2 are based on the researchers' assessment of the high-frequency words in each topic.

5.3 Outlook

The findings of the thesis point at additional research opportunities regarding the two main streams of literature. For example, future research might examine organized stakeholder responses and the (financial) consequences for firms, e.g., based on customer boycotts resulting from irresponsible corporate behavior with regard to the environment. This "stakeholder activism" might trigger costly firm actions to avoid "punishment" from the stakeholders. To examine the stakeholders' use of and reaction to non-financial information, the initiation of mandatory non-financial disclosure in the EU offers a suitable setting. This regulation requires large firms (>500 employees) to provide information about various non-financial aspects, for instance, product responsibility and the protection of human rights. Thus, the regulation might enable researchers to examine stakeholder groups, such as customers, to establish how the information demand of the various stakeholder groups might differ and what kind of consequences conflicts with the stakeholder groups might have for the firms. One challenge in this setting is to distinguish between CSR disclosure and CSR activities initiated in anticipation of the new disclosure requirements (e.g., Fiechter et al. (2019)). Nevertheless, the variation between the various European jurisdictions regarding the effective date and the disclosure threshold (less than 500 employees) allows researchers to employ research design choices suitable to establish causal inferences.

The lack of causal inferences also affects the literature about the financial press as an information intermediary. There are few papers (e.g., Engelberg and Parsons (2011)) that can link the market outcome directly to newspaper coverage. Therefore, future research should focus on settings allowing causal inferences such as the merger of large news organizations or external regulation of press outputs (e.g., the EU copyright directive for news aggregators and online media). Additionally, more research is needed on the difference between information production and dissemination outside the context of earnings announcements to paint a broader picture of the work of the press as an information intermediary. The extension of textual analysis, e.g., sentiment analysis or topic models, in accounting research might help to identify information production in other contexts (e.g., Loughran and McDonald (2016)).

A development that affects the general dissemination of information to investors and stakeholders is the extension of digital offerings and especially the rise of social media. For example, by disseminating their earnings announcements through Twitter, firms can effectively create the same attention as a newsflash. To that end, Blankespoor et al. (2014) show that Twitter dissemination lowers information asymmetries. If firms can substitute the dissemination role of the press using social media, the press must provide other services such as information production to justify its position in the market. To save resources and focus on firms with current issues, the news outlets might also rely on computer-generated articles for less interesting firms (Blankespoor et al. (2018)). Therefore, it might be of interest for future research to examine how the information production changed over time considering the technological changes.

The dissemination of information via social media makes it easier for stakeholders to find it. Further, these platforms facilitate direct communication between the firms and the stakeholders, as most of them are two-way disclosure channels. Therefore, the stakeholders can raise issues more easily and organize adverse stakeholder reactions when the firms do not meet the demand of their stakeholders (e.g., Cade (2018)). Negative stakeholder feedback via social media might also imply reputational costs for the firms. Overall, the evolving disclosure landscape for firms and stakeholders, as well as technological innovations, offer a variety of fruitful avenues for future research about the dissemination and production of financial and non-financial information.

6. Erklärung über den geleisteten Eigenanteil an der Arbeit

Study 1: *"Supply-side evidence on the role of the financial press as an intermediary of accounting information"*

- This study is single-authored.

Study 2: *"Information production by the financial press: A closer look"*

- This study was conducted in cooperation with Jörg-Markus Hitz (Georg-August University Göttingen) and Harm H. Schütt (Tilburg University). I was continuously involved and took part in the development of the concept of this study, the theoretical framework, and the writing of the text document. In particular, I was solely responsible for the data collection and the execution of the empirical analyses. I was partly responsible for the textual analysis elements included in the paper.

Study 3: *"Corporate social responsibility and stakeholder attention"*

- This study is single-authored.

Ann-Kristin Großkopf

7. Versicherung

gemäß §12 der Prüfungs- und Studienordnung (Version AM I 56/28.10.2016) für den Promotionsstudiengang „Wirtschaftswissenschaften" der Georg-August-Universität Göttingen.

Ich versichere,

1. dass ich die eingereichte Dissertation „Essays on the Media's Production and Dissemination Role: Evidence from financial and non-financial disclosure" selbstständig angefertigt habe und nicht die Hilfe Dritter in einer dem Prüfungsrecht und wissenschaftlicher Redlichkeit widersprechenden Weise in Anspruch genommen habe,
2. dass ich das Prüfungsrecht einschließlich der wissenschaftlichen Redlichkeit – hierzu gehört die strikte Beachtung des Zitiergebots, so dass die Übernahme fremden Gedankenguts in der Dissertation deutlich gekennzeichnet ist – beachtet habe,
3. dass beim vorliegenden Promotionsverfahren kein Vermittler gegen Entgelt eingeschaltet worden ist sowie im Zusammenhang mit dem Promotionsverfahren und seiner Vorbereitung
   - kein Entgelt gezahlt oder entgeltgleiche Leistungen erbracht worden sind
   - keine Dienste unentgeltlich in Anspruch genommen wurden, die dem Sinn und Zweck eines Prüfungsverfahrens widersprechen
4. dass ich eine entsprechende Promotion nicht anderweitig beantragt und hierbei die eingereichte Dissertation oder Teile daraus vorgelegt habe.

Mir ist bekannt, dass Unwahrheiten hinsichtlich der vorstehenden Versicherung die Zulassung zur Promotionsprüfung ausschließen und im Falle eines späteren Bekanntwerdens die Promotionsprüfung für ungültig erklärt werden oder der Doktorgrad aberkannt werden kann.

Göttingen, den 02.03.2020

Ann-Kristin Großkopf

References

Ahern, K. R. and Sosyura, D. (2015): Rumor Has It: Sensationalism in Financial Media. *Review of Financial Studies* 28(7): 2050-2093.

Ahern, Kenneth R. and Sosyura, Denis (2014): Who Writes the News? Corporate Press Releases during Merger Negotiations. *The Journal of Finance* 69(1): 241-291.

Armstrong, Christopher S., Barth, Mary E., Jagolinzer, Alan D., and Riedl, Edward J. (2010): Market Reaction to the Adoption of IFRS in Europe. *The Accounting Review* 85(1): 31-61.

Barber, Brad M. and Odean, Terrance (2008): All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors. *Review of Financial Studies* 21(2): 785-818.

Basu, Sudipta, Duong, Truong Xuan, Markov, Stanimir, and Tan, Eng-Joo (2013): How Important are Earnings Announcements as an Information Source?, *European Accounting Review* 22(2): 221-256.

Beaver, William H. (1998). *Financial Reporting – An Accounting Revolution*. Upper Saddle River.

Bednar, M. K. (2012): Watchdog or Lapdog? A Behavioral View of the Media as a Corporate Governance Mechanism. *Academy of Management Journal* 55(1): 131-150.

Ben-Nasr, Hamdi and Ghouma, Hatem (2018): Employee welfare and stock price crash risk. *Journal of Corporate Finance* 48: 700-725.

Bhattacharya, Nilabhra, Black, Ervin L., Christensen, Theodore E., and Mergenthaler, Richard D. (2007): Who Trades on Pro Forma Earnings Information?, *The Accounting Review* 82(3): 581-619.

Blankespoor, Elizabeth (2019): The Impact of Information Processing Costs on Firm Disclosure Choice: Evidence from the XBRL Mandate. *Journal of Accounting Research* 57(4): 919-967.

Blankespoor, Elizabeth, deHaan, Ed, and Zhu, Christina (2018): Capital market effects of media synthesis and dissemination: evidence from robo-journalism. *Review of Accounting Studies* 23(1): 1-36.

Blankespoor, Elizabeth, Miller, Gregory S., and White, Hal D. (2014): The Role of Dissemination in Market Liquidity: Evidence from Firms' Use of Twitter™. *The Accounting Review* 89(1): 79-112.

Blei, David M., Ng, Andrw Y., and Jordan, Michael I. (2003): Latent Dirichlet Allocation. *Journal of Machine Learning Research* 3: 993-1022.

Brown, Nerissa C., Crowley, Richard M., and Elliott, W. Brooke (2020): What Are You Saying? Using topic to Detect Financial Misreporting. *Journal of Accounting Research*: Forthcoming.

Bunderson, J. Stuart and Thompson, Jeffery A. (2009): The Call of the Wild: Zookeepers, Callings, and the Double-edged Sword of Deeply Meaningful Work. *Administrative Science Quarterly* 54(1): 32-57.

Bushee, Brian J., Core, John E., Guay, Wayne, and Hamm, Sophia J. W. (2010): The Role of the Business Press as an Information Intermediary. *Journal of Accounting Research* 48(1): 1-19.

Bushee, Brian J., Matsumoto, Dawn A., and Miller, Gregory S. (2003): Open versus closed conference calls: the determinants and effects of broadening access to disclosure. *Journal of Accounting and Economics* 34(1): 149-180.

Butler, Marty, Kraft, Arthur, and Weiss, Ira S. (2007): The effect of reporting frequency on the timeliness of earnings: The cases of voluntary and mandatory interim reports. *Journal of Accounting and Economics* 43(2-3): 181-217.

Cade, Nicole L. (2018): Corporate social media: How two-way disclosure channels influence investors. *Accounting, Organizations and Society* 68-69: 63-79.

Cahan, Steven F., Chen, Chen, Chen, Li, and Nguyen, Nhut H. (2015): Corporate social responsibility and media coverage. *Journal of Banking & Finance* 59: 409-422.

Cahan, Steven F., De Villiers, Charl, Jeter, Debra C., Naiker, Vic, and Van Staden, Chris J. (2015): Are CSR Disclosures Value Relevant? Cross-Country Evidence. *European Accounting Review* 25(3): 579-611.

Call, Andrew C., Emett, Scott A. , Maksymov, Eldar, and Sharp, Nathan (2018): Meet the Press: Survey Evidence on Financial Journalists as Information Intermediaries. *Available at SSRN: https://ssrn.com/abstract=3279453.*

Chen, Yi-Chun, Hung, Mingyi, and Wang, Yongxiang (2018): The effect of mandatory CSR disclosure on firm profitability and social externalities: Evidence from China. *Journal of Accounting and Economics* 65(1): 169-190.

Christensen, Dane M. (2016): Corporate Accountability Reporting and High-Profile Misconduct. *The Accounting Review* 91(2): 377-399.

Christensen, Hans B., Floyd, Eric, Liu, Lisa Yao, and Maffett, Mark (2017): The real effects of mandated information on social responsibility in financial reports: Evidence from mine-safety records. *Journal of Accounting and Economics* 64(2-3): 284-304.

Christensen, Hans B., Hail, Luzi, and Leuz, Christian (2019): Economic Analysis of Widespread Adoption of CSR and Sustainability Reporting Standards. *Available at SSRN: https://ssrn.com/abstract=3315673.*

Clement, Michael B. (1999): Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter?, *Journal of Accounting and Economics* 27(3): 285-303.

Core, John E., Guay, Wayne, and Larcker, David F. (2008): The power of the pen and executive compensation. *Journal of Financial Economics* 88(1): 1-25.

Coyne, Joshua, Kim, Kevin H., and Kim, Sangwan (2019): The Role of the Business Press in Creating and Disseminating Information around Earnings Announcements. *European Accounting Review*: 1-29.

Da, Zhi, Engelberg, Joseph, and Gao, Pengjie (2011): In Search of Attention. *The Journal of Finance* 66(5): 1461-1499.

Dai, Lili, Parwada, Jerry T., and Zhang, Bohui (2015): The Governance Effect of the Media's News Dissemination Role: Evidence from Insider Trading. *Journal of Accounting Research* 53(2): 331-366.

deHaan, Ed, Lawrence, Alastair, and Litjens, Robin (2019): Measurement Error in Dependent Variables in Accounting: Illustrations Using Google Ticker Search and Simulations. *Available at SSRN: https://ssrn.com/abstract=3398287.*

Dhaliwal, Dan S., Li, Oliver Zhen, Tsang, Albert, and Yang, Yong George (2011): Voluntary Nonfinancial Disclosure and the Cost of Equity Capital: The Initiation of Corporate Social Responsibility Reporting. *The Accounting Review* 86(1): 59-100.

Dhaliwal, Dan S., Radhakrishnan, Suresh, Tsang, Albert, and Yang, Yong George (2012): Nonfinancial Disclosure and Analyst Forecast Accuracy: International Evidence on Corporate Social Responsibility Disclosure. *The Accounting Review* 87(3): 723-759.

Dimson, Elroy, Karakaş, Oğuzhan, and Li, Xi (2015): Active Ownership. *Review of Financial Studies* 28(12): 3225-3268.

Dougal, Casey, Engelberg, Joseph, García, Diego, and Parsons, Christopher A. (2012): Journalists and the Stock Market. *Review of Financial Studies* 25(3): 639-679.

Drake, Michael S., Guest, Nicholas M., and Twedt, Brady J. (2014): The Media and Mispricing: The Role of the Business Press in the Pricing of Accounting Information. *The Accounting Review* 89(5): 1673-1701.

Drake, Michael S., Jennings, Jared, Roulstone, Darren T., and Thornock, Jacob R. (2017): The Comovement of Investor Attention. *Management Science* 63(9): 2847-2867.

Drake, Michael S., Roulstone, Darren T., and Thornock, Jacob R. (2012): Investor Information Demand: Evidence from Google Searches Around Earnings Announcements. *Journal of Accounting Research* 50(4): 1001-1040.

Dyck, Aleander and Zingales, Luigi (2003): The Media and Asset Prices Working Paper.

Dyck, Alexander, Volchkova, Natalya, and Zingales, Luigi (2008): The Corporate Governance Role of the Media: Evidence from Russia. *The Journal of Finance* 63(3): 1093-1135.

Dyer, Travis, Lang, Mark, and Stice-Lawrence, Lorien (2017): The evolution of 10-K textual disclosure: Evidence from Latent Dirichlet Allocation. *Journal of Accounting and Economics* 64(2-3): 221-245.

Edmans, Alex (2011): Does the stock market fully value intangibles? Employee satisfaction and equity prices. *Journal of Financial Economics* 101(3): 621-640.

Edmans, Alex (2012): The Link Between Job Satisfaction and Firm Value, With Implications for Corporate Social Responsibility. *Academy of Management Perspectives* 26(4): 1-19.

Eickhoff, Matthias and Neuss, Nicole (2017): Topic Modelling Methodology: Its Use In Information Systems And Other Managerial Disciplines. 25th European Conference on Information Systems (ECIS), Guimarães, Portugal, June 5-10, 2017, http://aisel.aisnet.org/ecis2017_rp/86.

Elliott, W. Brooke, Hobson, Jessen L., and Jackson, Kevin E. (2011): Disaggregating Management Forecasts to Reduce Investors' Susceptibility to Earnings Fixation. *The Accounting Review* 86(1): 185-208.

Engelberg, Joseph and Parsons, Christopher A. (2011): The Causal Impact of Media in Financial Markets. *The Journal of Finance* 66(1): 67-97.

Faleye, Olubunmi and Trahan, Emery A. (2011): Labor-Friendly Corporate Practices: Is What is Good for Employees Good for Shareholders?, *Journal of Business Ethics* 101(1): 1-27.

Fang, Lily H., Peress, Joel, and Zheng, Lu (2014): Does Media Coverage of Stocks Affect Mutual Funds' Trading and Performance?, *Review of Financial Studies* 27(12): 3441-3466.

Fang, Lily and Peress, Joel (2009): Media Coverage and the Cross-section of Stock Returns. *The Journal of Finance* 5(64): 2023-2052.

Fiechter, Peter, Hitz, Joerg-Markus, and Lehmann, Nico (2019): Real Effects in Anticipation of Mandatory Disclosures: Evidence from the European Union's CSR Directive. *Available at SSRN: https://ssrn.com/abstract=3033883.*

Flammer, Caroline (2013): Corporate Social Responsibility and Shareholder Reaction: The Environmental Awareness of Investors. *Academy of Management Journal* 56(3): 758-781.

Foster, George (1979): Briloff and the Capital Market. *Journal of Accounting Research* 17(1): 262-274.

Freeman, R. Edward (1984). *Strategic management : a stakeholder approach*. Boston, Pitman.

Garrett, Jace, Hoitash, Rani, and Prawitt, Douglas F. (2014): Trust and Financial Reporting Quality. *Journal of Accounting Research* 52(5): 1087-1125.

Gartenberg, Claudine, Prat, Andrea, and Serafeim, George (2019): Corporate Purpose and Financial Performance. *Organization Science* 30(1): 1-18.

Gartenberg, Claudine and Serafeim, George (2019): Corporate Purpose and Firm Ownership. *Available at SSRN: https://ssrn.com/abstract=3440281*.

Glaeser, Stephen and Guay, Wayne R. (2017): Identification and generalizability in accounting research: A discussion of Christensen, Floyd, Liu, and Maffett (2017). *Journal of Accounting and Economics* 64(2-3): 305-312.

Goldman, Eitan, Martel, Jordan, and Schneemeier, Jan (2019): A Theory of Financial Media. *Available at SSRN: https://ssrn.com/abstract=3457591.*

Graf-Vlachy, Lorenz, Oliver, Abbie Griffith, Banfield, Richard, König, Andreas, and Bundy, Jonathan (2019): Media Coverage of Firms: Background, Integration, and Directions for Future Research. *Journal of Management* 46(1): 36-69.

Greening, Daniel W. and Turban, Daniel B. (2000): Corporate Social Performance As a Competitive Advantage in Attracting a Quality Workforce. *Business & Society* 39(3): 254-280.

Grewal, Jody, Riedl, Edward J., and Serafeim, George (2019): Market Reaction to Mandatory Nonfinancial Disclosure. *Management Science* 65(7): 3061-3084.

Guay, Wayne, Samuels, Delphine, and Taylor, Daniel (2016): Guiding through the Fog: Financial statement complexity and voluntary disclosure. *Journal of Accounting and Economics* 62(2-3): 234-269.

Guest, Nicholas M. (2018): Do Journalists Help Investors Analyze Firms' Earnings News? *Available at SSRN: https://ssrn.com/abstract=3067387*.

Guest, Nicholas M. and Jaewoo, Kim (2019): On the Interplay among Information Intermediaries: How Does Analyst Coverage Affect Media Coverage? *Available at SSRN: https://ssrn.com/abstract=3337116.*

Guiso, Luigi, Sapienza, Paola, and Zingales, Luigi (2015): The value of corporate culture. *Journal of Financial Economics* 117(1): 60-76.

Gurun, Umit G. and Butler, Alexander W. (2012): Don't Believe the Hype: Local Media Slant, Local Advertising, and Firm Value. *The Journal of Finance* 67(2): 561-597.

Hahn, Rüdiger and Kühnen, Michael (2013): Determinants of sustainability reporting: a review of results, trends, theory, and opportunities in an expanding field of research. *Journal of Cleaner Production* 59: 5-21.

Hawn, Olga, Chatterji, Aaron K., and Mitchell, Will (2018): Do investors actually value sustainability? New evidence from investor reactions to the Dow Jones Sustainability Index (DJSI). *Strategic Management Journal* 39(4): 949-976.

Healy, Paul M. and Palepu, Krishna G. (2001): Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of Accounting and Economics* 31(1): 405-440.

Heckman, James J. (1979): Sample Selection Bias as a Specification Error. *Econometrica* 47(1): 153-161.

Heese, Jonas and Perez-Cavazos, Gerardo (2020): When the Boss Comes to Town: The Effects of Headquarters' Visits on Facility-Level Misconduct. *The Accounting Review* forthcoming.

Hillert, Alexander, Jacobs, Heiko, and Müller, Sebastian (2014): Media Makes Momentum. *Review of Financial Studies* 27(12): 3467-3501.

Hitz, Joerg-Markus (2010): Press Release Disclosure of 'Pro Forma' Earnings Metrics by Large German Corporations – Empirical Evidence and Regulatory Recommendations. *Accounting in Europe* 7(1): 63-86.

Hoberg, G. and Maksimovic, V. (2014): Redefining Financial Constraints: A Text-Based Analysis. *Review of Financial Studies* 28(5): 1312-1352.

Hombach, Katharina and Sellhorn, Thorsten (2018): Shaping Corporate Actions Through Targeted Transparency Regulation: A Framework and Review of Extant Evidence. *Schmalenbach Business Review* 71(2): 137-168.

Hombach, Katharina and Sellhorn, Thorsten (2019): Firm Value Effects of Targeted Disclosure Regulation: The Role of Reputational Costs. *Available at SSRN: https://ssrn.com/abstract=3204505.*

Hope, Ole-Kristian, Hu, Danqi, and Lu, Hai (2016): The benefits of specific risk-factor disclosures. *Review of Accounting Studies* 21(4): 1005-1045.

Huang, Allen H., Lehavy, Reuven, Zang, Amy Y., and Zheng, Rong (2017): Analyst Information Discovery and Interpretation Roles: A Topic Modeling Approach. *Management Science*: 1-23.

Huang, Xuan, Teoh, Siew Hong, and Zhang, Yinglei (2014): Tone Management. *The Accounting Review* 89(3): 1083-1113.

Jensen, Michael C. (1979). Toward a Theory of the Press. Economics and Social Institutions. K. Brunner. Martinus Nijhoff Publishing Company**:** 267-287.

Jones, Thomas M. (1995): Instrumental stakeholder theory: a synthesis of ethics and economics. *Academy of Management Review* 20(2): 404-437.

Joos, Philip P. M. and Leung, Edith (2013): Investor Perceptions of Potential IFRS Adoption in the United States. *The Accounting Review* 88(2): 577-609.

Kothari, S. P., Li, Xu, and Short, James E. (2009): The Effect of Disclosures by Management, Analysts, and Business Press on Cost of Capital, Return Volatility, and Analyst Forecasts: A Study Using Content Analysis. *The Accounting Review* 84(5): 1639-1670.

Krüger, Philipp (2015): Corporate goodness and shareholder wealth. *Journal of Financial Economics* 115(2): 304-329.

Kuhnen, Camelia M. and Niessen, Alexandra (2012): Public Opinion and Executive Compensation. *Management Science* 58(7): 1249-1272.

Lennox, Clive S., Francis, Jere R., and Wang, Zitian (2012): Selection Models in Accounting Research. *The Accounting Review* 87(2): 589-616.

Leung, Edith and Veenman, David (2018): Non-GAAP Earnings Disclosure in Loss Firms. *Journal of Accounting Research* 56(4): 1083-1137.

Li, Congcong (2014): The hidden face of the media: How financial journalists produce information. *University of Maryland*: 64.

Li, Feng (2008): Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics* 45(2-3): 221-247.

Loughran, Tim and McDonald, Bill (2016): Textual Analysis in Accounting and Finance: A Survey. *Journal of Accounting Research* 54(4): 1187-1230.

Loughran, Tim and McDonald, Bill (2011): When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *The Journal of Finance* 66(1): 35-65.

Lys, Thomas, Naughton, James P., and Wang, Clare (2015): Signaling through corporate accountability reporting. *Journal of Accounting and Economics* 60(1): 56-72.

Madsen, Joshua and Niessner, Marina (2019): Is Investor Attention for Sale? The Role of Advertising in Financial Markets. *Journal of Accounting Research* 57(3): 763-795.

Madsen, Peter M. and Rodgers, Zachariah J. (2015): Looking good by doing good: The antecedents and consequences of stakeholder attention to corporate disaster relief. *Strategic Management Journal* 36(5): 776-794.

Matsumoto, Dawn, Pronk, Maarten, and Roelofsen, Erik (2011): What Makes Conference Calls Useful? The Information Content of Managers' Presentations and Analysts' Discussion Sessions. *The Accounting Review* 86(4): 1383-1414.

Miller, Gregory S. (2006): The Press as a Watchdog for Accounting Fraud. *Journal of Accounting Research* 44(5): 1001-1033.

Miller, Gregory S. and Skinner, Douglas J. (2015): The Evolving Disclosure Landscape: How Changes in Technology, the Media, and Capital Markets Are Affecting Disclosure. *Journal of Accounting Research* 53(2): 221-239.

Peress, Joel (2014): The Media and the Diffusion of Information in Financial Markets: Evidence from Newspaper Strikes. *The Journal of Finance* 69(5): 2007-2043.

Raghunandan, Aneesh (2019): How Are Non-Financial and Financial Misconduct Related? *Available at SSRN: https://ssrn.com/abstract=3307594.*

Rees, Lynn, Sharp, Nathan, and Twedt, Brady (2015): Who's heard on the Street? Determinants and consequences of financial analyst coverage in the business press. *Review of Accounting Studies* 20(1): 173-209.

Schütt, Harm H. (2019): Measuring Segmentation in the Financial News Market. *Available at SSRN: https://ssrn.com/abstract=3051732.*

Shipman, Jonathan E., Swanquist, Quinn T., and Whited, Robert L. (2017): Propensity Score Matching in Accounting Research. *The Accounting Review* 92(1): 213-244.

Soltes, Eugene (2009): News Dissemination and the Impact of the Business Press.

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

Tetlock, Paul C. (2011): All the News That's Fit to Reprint: Do Investors React to Stale Information?, *Review of Financial Studies* 24(5): 1481-1512.

Tetlock, Paul C., Saar-Tsechansky, Maytal, and Macskassy, Sofus (2008): More Than Words: Quantifying Language to Measure Firms' Fundamentals. *The Journal of Finance* 63(3): 1437-1467.

Thompson, Robert B., Olsen, Chris, and Dietrich, J. Richard (1987): Attributes of News about Firms: An Analysis of Firm-Specific News Reported in the Wall Street Journal Index. *Journal of Accounting Research* 25(2): 245-274.

Tsileponis, Nikolaos, Stathopoulos, Konstantinos, and Walker, Martin (2020): Do Corporate Press Releases Drive Media Coverage? *Available at SSRN: https://ssrn.com/abstract=2740045.*

Tucker, Jennifer Wu (2010): Selection bias and econometric remedies in accounting and finance research. *Journal of Accounting Literature* 29: 31-58

Twedt, Brady (2016): Spreading the Word: Price Discovery and Newswire Dissemination of Management Earnings Guidance. *The Accounting Review* 91(1): 317-346.

Wood, Donna, J. (1995): Stakeholder mismatching: a theoretical problem in empirical research on corporate social performance. *The International Journal of Organizational Analysis* 3(3): 229-267.

Zyglidopoulos, Stelios C., Georgiadis, Andreas P., Carroll, Craig E., and Siegel, Donald S. (2012): Does media attention drive corporate social responsibility?, *Journal of Business Research* 65(11): 1622-1627.