6-22-2020

Institutional versus Crowdsourced Analyst Reports: Who Puts it in a Nutshell?

Matthias Palmer
University of Goettingen, matthias.palmer@uni-goettingen.de

Steffen Bankamp
University of Goettingen, steffen.bankamp@uni-goettingen.de

Jan Muntermann
University of Göttingen, muntermann@wiwi.uni-goettingen.de

Institutional versus Crowdsourced Analyst Reports: Who Puts it in a Nutshell?

Completed Research Paper

Matthias Palmer
University of Goettingen
Platz der Goettinger Sieben 5
matthias.palmer@uni-goettingen.de

Steffen Bankamp
University of Goettingen
Platz der Goettinger Sieben 5
steffen.bankamp@uni-goettingen.de

Jan Muntermann
University of Goettingen
Platz der Goettinger Sieben 5
muntermann@wiwi.uni-goettingen.de

Abstract

Crowdsourced reports of nonprofessional analysts published on online platforms enjoy increasing popularity. At the same time, it is reported that analyst reports of institutional equity research firms are becoming less influential. In this study, we aim to explain these two phenomena by comparing the information provision in texts of institutional and crowdsourced analyst reports. In our comparative analysis, we apply text mining techniques to evaluate and compare (i) readability and (ii) information density of more than 25,000 analyst reports. We find better readability for crowdsourced than for institutional analyst reports. Furthermore, the crowdsourced reports provide more information within the same length of text. With this study, we provide evidence for explaining the success of crowdsourced analyst reports. Based on these insights, we also provide established brokerage houses an indication of how they could improve their reports.

Keywords: Social media analysts, analyst reports, text readability, information density, crowdsourcing

Introduction

Recent research concludes that social media equity research is disrupting the work of traditional financial analysts (Drake et al. 2019). In particular, it is emphasized that there is an apparent decline in the scope and importance of sell-side equity research. The number of equity analysts employed at twelve major international banks has decreased by 20 % compared to 2012. After a sharp decline in revenues in recent years, research budgets for sell-side equity research are expected to fall by 20 % to 30 % in 2020 due to changes in regulations (Lee 2019). Fang et al. (2019) found decreased importance of sell-side reports, as investment companies tend to insource equity research. Sell-side equity research refers to research that is carried out by financial analysts who work for stockbrokers. Furthermore, recent research indicates that the output of *nonprofessional analysts* (NPA) can directly be associated with price changes in capital markets (Campbell et al. 2019), underlining their relevance and success. NPAs differ from *professional analysts* (PA) in that they are not regularly paid to prepare reports, they are not employed by any equity research firm or brokerage house, and that their reports are available online mostly free of charge to any investor. Since basically anyone can write these reports, we collectively refer to them by using the term *crowdsourced analyst reports*. By *institutional analyst reports*, we refer to reports of typical sell-side PAs. Usually, these reports require a fee to be paid or are distributed to brokerage house clients.

By analyzing the interplay between PAs and NPAs, Drake et al. (2019) find a reduced market reaction to forecasts by PAs in cases where reports by NPAs have been previously published. Furthermore, they report that given the same information availability, NPAs tend to publish new information prior to PAs. However, Drake et al. (2019) did not include textual data, although it is an essential output of both PAs and NPAs. Textual features are particularly relevant because concise writing lowers information-processing costs. Obfuscation theory provides an explanation of why texts are not written as concise as possible. Li (2008) uses this theory to explain why managers use opaque language in annual reports, e.g., to obfuscate negative news. Obfuscation can be either intended or unintended. The paper by Li (2008) assumes that obfuscation is intended in any case, since annual reports usually contain well-considered texts. For analyst reports, in contrast to De Franco et al. (2015), we take a somewhat more differentiated perspective. Due to various influences especially the texts of inexperienced analysts can show obfuscation. We see the potential for unintended obfuscation if unexperienced PAs are not able to express themselves more clearly or are under time pressure. Besides that, intended obfuscation could arise due to career pressure when ambitious PAs want to draw attention to themselves with a particularly extravagant language. For experienced analysts, time pressure can lead to unintentional obfuscation. For crowdsourced analyst reports, we consider the situation to be different. Of course, unintended obfuscation can also occur here if the NPAs are bad at writing texts concisely. Beyond that, we do not see any major risks due to time or career pressure. Therefore, we see a considerably higher risk of obfuscation for PAs than for NPAs. From our perspective, this also means that lower information processing costs can be expected for crowdsourced analyst reports. Overall, obfuscation leads to less understandable phrasing, which can be analyzed by readability measures (how easy is a text to understand) and information density (how much information is conveyed in a text).

For financial analysts, there is only literature on the readability of institutional analyst reports. In general, it has been shown that hard to read financial disclosures lead to smaller capital market reactions (Lawrence 2013; Rennekamp 2012). For annual reports, Li (2008) shows that the readability and also text length can be related to the earnings of the company. For example, Courtis (1986) calculate Fog and Flesch readability scores for annual reports and reveal that these texts are written inappropriately difficult for the target audience. In summary, for a range of comparable studies, this means that annual reports contain texts that are hard to read and are beyond the reading skills of 90 % of the adult population and even beyond the skills of 40 % of investors (Courtis 1995). Building on this, De Franco et al. (2015) explicitly deals with the work of PAs. They conclude that by increasing analyst reports' readability, they can be read and understood more quickly, thereby reducing information processing costs. Further, they show that this is in the interest of sell-side PAs, as they anticipate that this will lead to increased trading volume carried out through their brokerage services generating commission fees. De Franco et al. (2015) further find more readable reports published from so-called "high-ability" PAs (ranked by the Institutional Investor magazine). They relate easier to read reports to more experienced, more successful, and more active PAs, who adjust their buy recommendations and price targets more frequently. De Franco et al. (2015) measure a negative correlation between text readability and firm complexity. Interestingly, this study concludes that PAs' industry experience has a positive impact on their reports' readability. Hsieh et al. (2016) take a more capital market-oriented approach and examine the extent to which well-readable institutional analyst reports can reduce uncertainties in earnings expectations and thus lead to an increase in share prices. They conclude that the capital market reacts positively to institutional analyst reports that are more readable. Asay et al. (2016) find that when companies publish poorly readable disclosures, investors are increasingly turning to text that is non-firm information, e.g., analyst reports and news media. They arrive at this conclusion because they observe that investors feel less comfortable reading texts that are difficult to understand. In our view, this implies that it is crucial for analyst reports to be easy to read. And it is precisely here that we see a parallel to the relationship between institutional and crowdsourced analyst reports.

Obfuscation should result in less information density. Contrary to readability scores that have been frequently applied in the financial context, there is a lack of finance-related studies on information density. Therefore, we make use of findings from the field of density-based content extraction (Annam and Sajeev 2016). Here, Entropy is a useful tool to analyze information density (Dethlefs et al. 2012).

Against this background and the emergence of crowdsources analyst reports, we examine how well PAs and NPAs manage to write easy-to-read texts and to present information as concisely as possible. Overall, we aim to contribute to a better understanding of the growing success of crowdsourced analyst reports. Therefore, we ask the following research question:

RQ: *How does information provision differ between texts of institutional and crowdsourced analyst reports?*

In the following, we present the theoretical background on analyst reports and introduce text readability and information density scores. Then, we present the dataset and explain the preparation of the textual data. In the analysis section, we compare the readability and information density in institutional and crowdsourced analyst reports and additionally validate our results by alternative analytical approaches. The paper closes with a discussion and a synthesis of our findings.

Theoretical Background

Analyst Reports

Asquith et al. (2005) describe institutional analyst reports as the result of the collection, evaluation, and dissemination of information that might be relevant to the future of a company. The textual part of institutional analyst reports contains, for example, assessments of new, publicly available information, but also new information that has been researched by the PAs (Chen et al. 2010). Institutional analyst reports consist of three parts: A stock recommendation (often divided into strong buy, buy, hold, sell, and strong sell), a price target (usually for 6 or 12 months in the future), and earnings forecast (for the next fiscal quarter or an entire fiscal year) (Womack 1996). Asquith et al. (2005) indicate that half of their examined institutional analyst reports contain new information that was previously not publicly available. Furthermore, they find that investors especially use information from institutional analyst reports when PAs have downgraded their recommendations. Reports by PAs are particularly valuable because they also contain information derived from private communication between the PAs and the senior management of the companies (Brown et al. 2015). For example, Frankel et al. (2006) show that analyst reports are less informative when information processing costs are higher. This is in line with the fact that a particularly large number of institutional analyst reports are published after the release of quarterly earnings announcements (Huang et al. 2017).

Crowdsourced analyst reports are reports that are oftentimes freely available and therefore particularly interesting for private investors. NPAs write these reports generally without direct or indirect financial compensation. Therefore, it can be assumed that if social media analysts want their reports to be read frequently by many investors, their reports must meet high quality standards. Because these reports are provided at no charge, the information should be provided as concisely as possible and they should be relatively easy to read. In contrast to reports by institutional analysts, which are very similar in their structure (often predetermined by a given layout of a brokerage house) and contain textual parts on legal aspects, there are no such rigid specifications for crowdsourced analyst reports and only short disclaimers. Concerning legal aspects, according to Campbell et al. (2019), there is no negative impact on the credibility of the reports if NPAs hold stock positions in the company of the respective report. Nevertheless, Drake et al. (2019) give three reasons why the content of crowdsourced analyst reports should be treated with caution: First, it is difficult to assess the financial skills of NPAs, e.g., because they have not gone through an application process. Second, unlike NPAs, PAs are subject to certain compliance rules and are controlled by financial regulators. Third, NPAs are not financially dependent on the reports they write and therefore have no fear of suffering a loss of reputation or losing a job. This could result in them being less careful when preparing their reports.

Text Readability and Information Density

The readability of a text indicates for which target audience it is most suitable and how easy it is to understand. Text readability is mainly determined by the used vocabulary and sentence length. However, current readability measures do not account for text and sentence structure. The most commonly used readability scores include the Flesch Reading Ease Score (Flesch and Gould 1949), the Flesch Kincaid Grade (Kincaid et al. 1975), and the Gunning Fog Index (Gunning 1952). Following Štajner et al. (2012) we provide short explanations and the formulas for the scores.

The Flesch Reading Ease score takes the average sentence length and the average number of syllables of the document into account. This results in a score between 0 and 100. Lower scores are associated with an increase in the text difficulty. For example, a score between 30–50 is considered difficult and is associated with an attainment level of an undergraduate degree. The score is calculated as follows:

$$\text{Flesch Reading Ease} = 206.835 - \left(1{,}015 \ \frac{words}{sentences}\right) - \left(84{,}6 \ \frac{syllables}{words}\right) \tag{1}$$

The Flesch Kincaid readability formula is a modification of the Flesch Reading Ease score. The advantage of this formula is that it estimates readability according to the US education level (in grades) and is therefore easy to interpret. Higher scores indicate increased difficulty. Similar to the Flesch Reading Ease score, the formula is based on the average sentence length of a document and the average number of syllables per word in the text:

$$\text{Flesch Kincaid Readability} = 0.39 \ \frac{words}{sentences} + 11.8 \ \frac{syllables}{words} - 15.59 \tag{2}$$

The (Gunning) Fog Index indicates whether a text can be understood on first reading by a person who has a certain number of years of education. It analyzes the relationship between average sentence length and the (percentage) number of complex words (more than two syllables) for every 100 words in a text. The index increases with difficulty. Scores between 14 and 18 correspond to a high level of difficulty with the corresponding number of years of education. A score of 17 corresponds to a college graduate level. In several academic studies, the Fog Index is the measure of choice for texts with finance-related content (Loughran and McDonald 2014). Nevertheless, in addition to the Fog Index, we also apply the Flesch Reading Ease score and the Flesch Kincaid readability in our study, as we want to ensure that our results are consistent under different calculation methods. On the one hand, Loughran and McDonald (2014) conclude that the readability measures are not as suitable as assumed and propose document length as an alternative measure for finance-related texts. On the other hand, in our view, the document length is not suitable for our research approach since we already know that structurally strong differences exist between institutional and crowdsourced analyst reports. That would not be the difference this paper aims to examine. The Fog Index formula is as follows:

$$\text{Fog Index} = 0.4 \ \left(average\ sentence\ length + complex\ words\right) \tag{3}$$

Shannon Entropy (Shannon 1948) is used to assess the information density in a text. Entropy can be considered a quantification of the probability of a future event. For text mining applications, this corresponds to the uncertainty about the identity of a subsequent word in a sentence (next-word Entropy). Thus, if it is evident in a text which word must follow after a given part of the text (there is no uncertainty about the identity of the subsequent word), the Entropy is zero. If all possible words are equally likely to be the subsequent word, maximum Entropy is given. Aurnhammer and Frank (2019) define the next-word Entropy $H(t)$ of the distribution of a subsequent word W as follows:

$$H(t) = -\sum_{W_{t+1} \in W} P(W_{t+1} | W_{1\ldots t}) \log P\,(W_{t+1} | W_{1\ldots t}) \tag{4}$$

Given the theoretical explanations of the two different types of analyst reports and the measures of text readability and level of information, we propose four research hypotheses. First, we want to ensure that the two types of reports are suitable for comparison (**H1**). Then, we want to compare the readability of the report types with each other (**H2**). By manually reviewing the reports, we realized that particularly technical aspects are usually explained in detail at the end of the reports. We, therefore, want to check whether this is also reflected in the readability scores (**H3**). Finally, we need to compare the information comprehension of the two report types (**H4**) since we expect a relationship between readability and information density.

H1: *Chronologically related institutional and crowdsourced analyst reports discuss similar content regarding a specific company and are therefore substitutes and suitable for comparative analysis.*

H2: *The readability of institutional analyst reports is higher than the readability of crowdsourced analyst reports.*

H3: *For both institutional and crowdsourced analyst reports, the first half of the report is easier to read than the second half.*

H4: *In contrast to institutional analyst reports, crowdsourced analyst reports contain more information for the same text length.*

Dataset and Descriptive Statistics

We collect institutional analyst reports from Thomson Reuters as well as crowdsourced analyst reports from a major online platform providing crowdsourced equity research. The investigation period of four years ranges from 01-01-2015 to 12-31-2018. We chose the Dow Jones Industrial Average (DJIA) as a company sample. This index contains 30 of the largest companies in the United States. The constituents of the DJIA are particularly appropriate for our study as they receive high attention from analysts. All 32 companies that have been a constituent of the DJIA during the observation period are considered. Our data sample consists of 25,893 reports published by 1,092 different authors. For the crowdsourced analyst reports, the authors can easily be identified by their username. In the case of institutional reports, we refer to the brokerage house that publishes the report when we use the term author. Approximately, one quarter of the reports is crowdsourced and three quarters originate from institutional sources.

The crowdsourced reports were written by 928 unique authors and the larger sample of institutional reports was written by 164 authors. Table 2 in the appendix gives an overview of the dataset. The publication volume at the company level is strongly associated with the two report types. Apple is the most covered company of the sample, both for crowdsourced and institutional reports. When selecting the institutional analyst reports, we ensure that they relate to a single company only. Following De Franco et al. (2015), we exclude so-called "morning meeting notes" from the analyst reports dataset. These documents are not regular reports but short daily updates on the companies covered.

Figure 1 shows the publication pattern for both sources over time. A strong co-movement between both time series is observable. Within each year, four major peaks can be recognized. These peaks lay around the date of quarterly earnings figures. This suggests that both groups are exposed to similar information. Since the overall publication volume remains relatively stable in the years 2015 to 2017, a decline in crowdsourced reports can be observed in 2018.

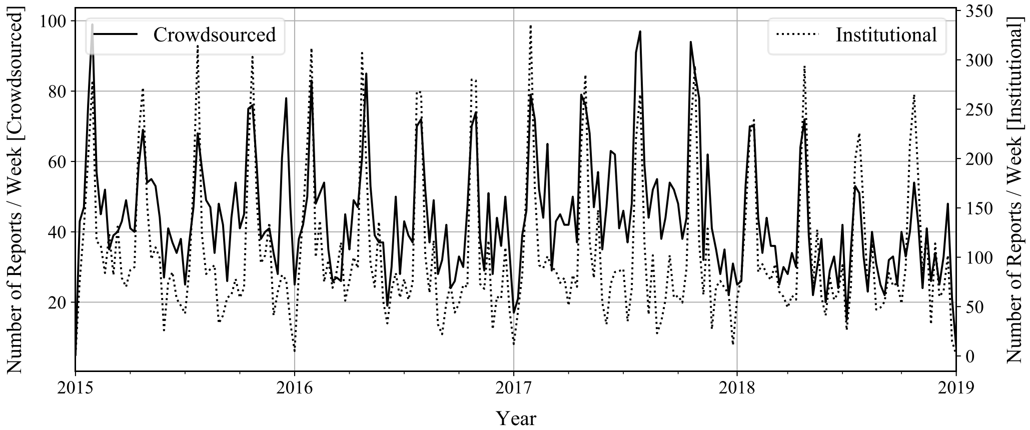

Figure 1. Publication pattern over time

After extracting the text from the documents, we follow Huang et al. (2017) and remove tables, figures (including captions), disclaimers, and boilerplate. The mean document length is 1,302 words for institutional and 1,128 words for crowdsourced analyst reports. However, the mean document length of institutional reports is strongly influenced by outliers since the longest document amounts to over 13,000 words. Considering the median, crowdsourced reports are longer (1,022 words in crowdsourced and 929 in institutional reports). We prepare the text in two different ways for our analyses. The calculation of readability scores requires completely preserved sentences, as stopword removal results in an underestimation of sentence length and stemming in an underestimation of word length. We edit the text in such a way that each sentence is separately identifiable. To calculate text similarity and Entropy, we carry out further pre-processing. We transform the text to lower case and tokenize it. We also remove stop words and stem the remaining words. Finally, we remove words that consist of two letters. The aforementioned steps were conducted using the Python packages *nltk* and *readability*.

Readability measures are essentially based on word and sentence lengths. To ensure the validity of the results, it is important to check the plausibility of these two measures. For this purpose, we first checked manually on a randomly selected sample whether the implemented pre-processing works as expected. We paid special attention to the correct sentence and word tokenization. Besides, we examine the distribution of word and sentence length over the document corpus for each document type (see Figure 2). Both sentence and word length have reasonable distributions. It is therefore unlikely that our results are driven by systematic errors within the dataset.

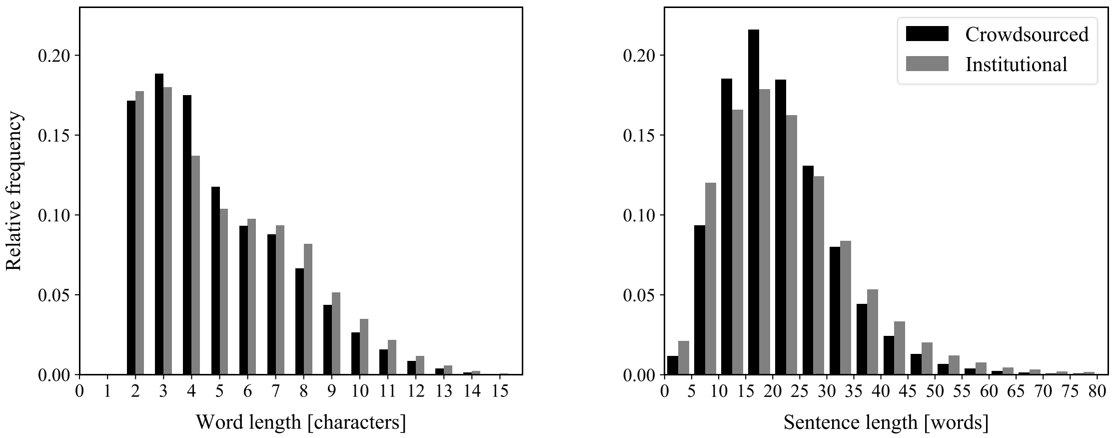

Figure 2. Distribution of word and sentence tokens by their length

We observe a clear pattern of institutional authors making more use of long words (six or more letters) than authors of crowdsourced reports (see Figure 2). Short words are more prevalent in crowdsourced reports. On the two-letter level, we see a deviation from that relationship. Further investigation reveals that this can be explained by the heavy use of abbreviations. Although the words "*of*", "*to*" and *"in"* are the most common two-letter words in both types of documents, abbreviations are also common. We labeled the 100 most used two-letter words of each document type as either normal word or abbreviation. Abbreviations that are common in linguistic usage have been classified as normal words (e.g., UK). We found 54 abbreviations in the institutional and 51 in the crowdsourced word list. When weighting these results according to their occurrence, we find 6.44 % of abbreviations in institutional reports compared to 1.62 % in their crowdsourced peers. This suggests a strong usage of professional jargon in institutional analyst reports. The most used abbreviation in both document types is "*FY"* (fiscal year). The usage of abbreviations in two-letter words is most likely to be higher, as we only labeled the top 100 words, where the actual two-letter words are overrepresented.

Similar to word tokens, we find an overrepresentation of either short (10 words or less) or very long sentences (more than 30 words) in reports written by PAs, whereas the NPAs are more stuck to sentences of medium length. The mean sentence length of reports from PAs amounts to 23.34 words versus 21.74 words for reports written by NPAs. The higher mean and particularly the large number of very long sentences in the reports of PAs indicate that their texts are more difficult to read. In conclusion, the different word and sentence lengths are measuring different types of formulation and are therefore suitable to measure readability. Furthermore, the results are not based on outliers.

Analysis

Similarity of Report Types

In our motivation of the paper, we have shown that a comparison of the two report types is of great interest from a scientific point of view. Nevertheless, we must ensure that the two document types cover the same content. If this is not the case, the knowledge gained from the comparison of their readability would be of little importance. The documents are comparable if they are substitutes. Here, we refer to **H1** as the substitution hypothesis. Jame et al. (2017) conclude that institutional analysts adjust their actions in case of increased coverage of the respective company through crowdsourced reports. This behavior can only be rationally justified if institutional investors believe in the substitution hypothesis.

H1 can additionally be supported by our data. The publication volume of both document types is very synchronous, as shown in Figure 1. The pattern can either be explained by authors reacting to identical events of the followed companies or by similar information demands of their readers. We have also explored this hypothesis at the textual level by calculating the cosine similarity from the term-document-matrix between all report pairs of the same company (see Figure 3).

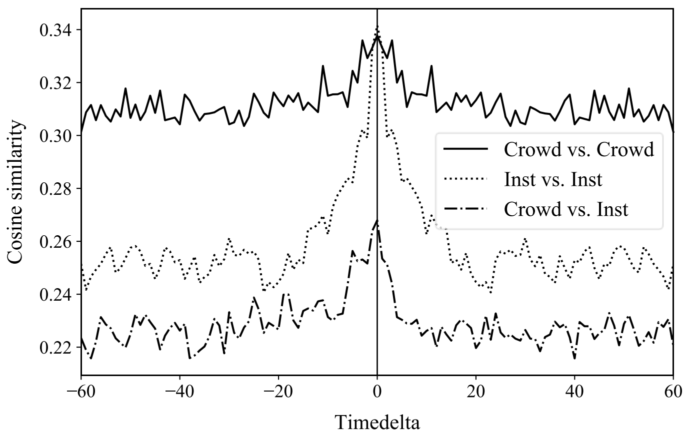

Figure 3. Cosine similarity between report pairs

The solid line and the dotted line are axisymmetric to the vertical line at day zero as the reports are matched in both directions. The dash-dotted line is not symmetric as the matching is conducted from crowdsourced to institutional reports only. In this case, a positive time delta means that the crowdsourced report was published before the paired institutional report.

In Figure 3, we compared both the reports within a report type as well as between the two report types. The mean of the cosine similarities is plotted against the time difference of the document pairs. We observe the highest similarity between the crowdsourced reports. The similarity between institutional analyst reports is lower. Only for day 0, we see a higher similarity of institutional reports. The overall lowest similarity is observed between mixed pairs (institutional and crowdsourced). However, the key message of this plot is that all three lines are peaking when the time difference becomes zero. Applying Welch's t-test shows (for all three pairing groups shown in Figure 3) that reports published within 5 days have significantly ($p < 0.001$) higher similarity than pairs published with the higher time difference. As these findings might be less surprising for intra-group similarity, the similarity between institutional and crowdsourced reports reveals an interesting insight into the relation of the document types. It can be interpreted that both authors are exposed to similar information and that they incorporate this information into their reports. This also provides evidence for the substitution hypothesis. The sharp increase of the dash-dotted line (Crowd vs. Inst) at day −7 combined with the steep decline at day +2 indicates that the PAs are rather ahead of the NPAs. In summary, the aforementioned arguments provide strong evidence for **H1**. Therefore, a comparison of the two document types seems possible.

Report Readability

Before evaluating readability, an analysis on a word-by-word basis is the first step. The mean average word length of institutional reports is 6.25 characters versus 5.98 for crowdsourced reports. Figure 4 shows the distribution of complex word usage per report. According to the Fog Index, a word is defined as complex if it consists of three or more syllables. On average, the institutional reports have a significantly ($p < 0.001$) higher proportion of complex words (19.42 %) than crowdsourced reports (16.26 %). Since text with less complex words is assumed to be easier to read (Li 2008), this finding supports the hypothesis **H2** of better readability of crowdsourced analyst reports. The sentence length is the second parameter incorporated in the most common readability measures. Long sentences are associated with more difficult readability. Similar to the word-by-word analysis, we calculate average sentence lengths by dividing the number of words of each document by the number of sentences. The mean average sentence length of institutional reports is significantly ($p < 0.001$) higher (23.34 words) compared to crowdsourced reports (21.74). Regarding hypothesis **H2**, both findings for word and sentence level suggest easier readability of crowdsourced reports.

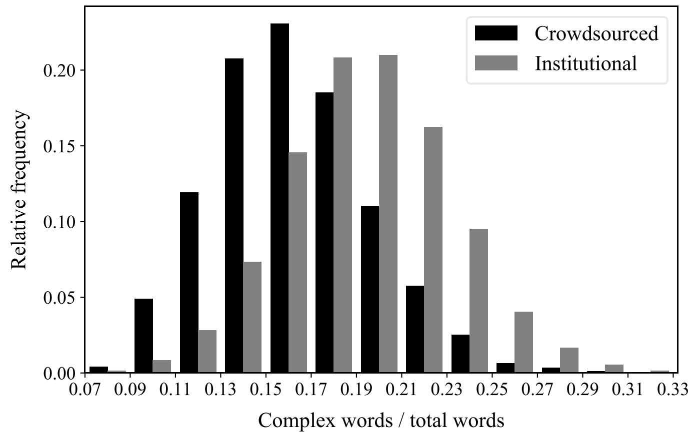

Figure 4. Word complexity by document type

To combine the two dimensions of readability, we use established readability measures. As stated in Table 1, we can observe a significant difference between the two document types across different readability scores. For institutional analyst reports, we measure a mean Fog Index (3) of 17.10, which exceeds the Fog Index of crowdsourced reports (15.20) substantially. The magnitude of the measured Fog Index is also consistent with previous studies. De Franco et al. (2015) found a mean Fog Index of 18.71 in institutional analyst reports. A study by Hsieh et al. (2016) measures a noticeable lower Fog Index of 14.01. The value we measured lies within the range of these two studies and thus seems plausible. Also, our value is below what was measured as readability for annual reports. These reports are published by companies and are a major information source for analysts. In this kind of financial documents, Loughran and McDonald (2014) find a Fog Index of 18.68 and Li (2008) an even higher Fog Index of 19.39. Since it is the task of analysts to prepare and aggregate information, it is not surprising that the readability score of their reports is lower compared to annual reports. To the best of our knowledge, we are the first to analyze the readability of crowdsourced reports. Therefore, we cannot meaningfully relate the magnitude of the Fog Index to findings of other studies.

Table 1. Welch's t-test of differences in mean readability scores

Readability Measures	Mean		STD		Δ Mean	t-value
	Crowd.	Inst.	Crowd.	Inst.	Crowd. − Inst.	
Fog Index	15.20	17.10	2.15	3.00	−1.90***	−55.50
Flesch Reading Ease	60.64	55.50	9.51	13.56	5.14***	33.58
Flesch Kincaid Grade	10.20	11.32	2.03	3.02	−1.11***	−33.55

*** $p < 0.001$

For the Flesch Reading Ease (1), we get a positive difference (see Table 1), as a higher score is associated with the easier-to-read text. For the Flesch Kincaid Grade (2), we obtain a negative difference in means, similar to the Fog Index. Our findings are robust and independent of the selection of the readability measure. To illustrate the difference between the report types, the distributions are plotted in Figure 5. The results support **H2:** Crowdsourced analyst reports provide content that is easier to read.

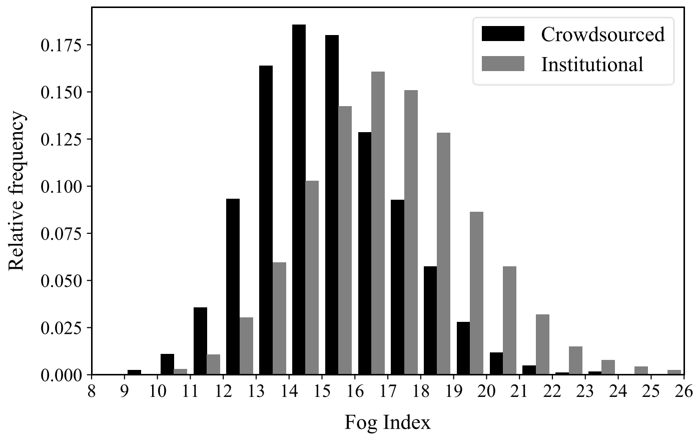

Figure 5. Readability score by document type

We showed that abbreviations are major text components of institutional reports. But how does the usage of abbreviations instead of fully written words effects measured readability? Abbreviations will not impact the number of sentences of a document. The word number, however, is likely to decrease, as word compositions like "*fiscal year*" are shortened to "*FY*" which is recognized as a single word by the tokenizer. The same applies to the number of syllables. If a word with more than two syllables is abbreviated to a two-letter word, it will reduce the number of complex words within the document. Abbreviations can reduce sentence length and the proportion of complex words. As both are incorporated in the Fog Index, the use of abbreviations will improve the measured readability of a text. Replacing the abbreviations by their original word would increase the Fog Index. However, due to the stronger usage of abbreviations in institutional reports, the increase would be much larger here. Also, the difference in readability would even be larger. Therefore, our findings are robust against the use of abbreviations. The results are not a consequence of an uneven usage of abbreviations across the document types.

The evidence for hypothesis **H2** could be due to an omitted variable problem rather than a difference in the formulation habits of the different analyst groups. Different coverage of the individual companies by the two analyst groups might have an impact here. For example, NPAs could especially cover companies with easy-to-understand B2C business models, while PA could also focus on companies with a more complex B2B focus. In such a case, we would not measure the difference between the analyst groups but the difference between the companies covered. The same reasoning can be applied to the publication pattern. For example, the NPAs write a disproportionately large number of reports between quarterly releases compared to the PAs (see Figure 1). This report type could be simpler in its basic structure. To exclude these effects, we apply a report pairing. We build pairs based on the company covered and the publication date. Pairs are only built if the time difference is no longer than ten days. Each report is matched once. We build 5,431 pairs, whereas 26.37 % of the pairs are published on the same day. This sample is robust against effects from publication patterns and coverage selection. This adjusted analysis confirms our previous results on **H2** without exception. The highly significant excess readability score of institutional reports amounts to 2.00 for the Fog Index.

Bisection of Reports

In the previous analyses, we utilized entire documents to calculate readability measures. However, the question arises whether it is justified to compare entire documents and thus giving equal weight to all textual components. Similar to scientific papers, the reader can read a document partially. In the case of scientific papers, the abstract would certainly be of particular relevance. This can also be applied to analyst reports. In this case, we assume that the author places the essential contents at the beginning and more technical passages at the end of the document. To control for an uneven distribution of readability over the document, we follow a straight-forward approach and split the documents into two parts. A breakdown based on report sections would be desirable, but the sample size does not allow a manual separation. Also, the reports are too heterogeneous to perform the breakdown automatically.

As we can learn from Figure 6, the difference in readability between the two report types is smaller for the first half of the documents than for the second half. In the first half, the excess Fog Index score of PA amounts to 1.67 versus 2.31 in the second half. Both differences are highly significant ($p < 0.001$). The findings from the comparison of the entire documents also apply to their bisection. However, this analysis provides further insights into the structure of crowdsourced and institutional analyst reports. The two plots in the lower part of Figure 6 compare the bisection within each report type and provide explanations for the effect we see in the upper part. Both document types are becoming harder to read when the reader reaches the rear section. We conclude that the authors provide the main facts of their analysis in a condensed and easier to understand format in the first part of their reports. The more complex and in-detail explanations are written in the last sections where the authors might be less concerned with readability. We find highly significant ($p < 0.001$) negative first half excess readability, which supports **H3**. However, the size of the effect is stronger for PA (−0.85) compared to NPA (−0.21).

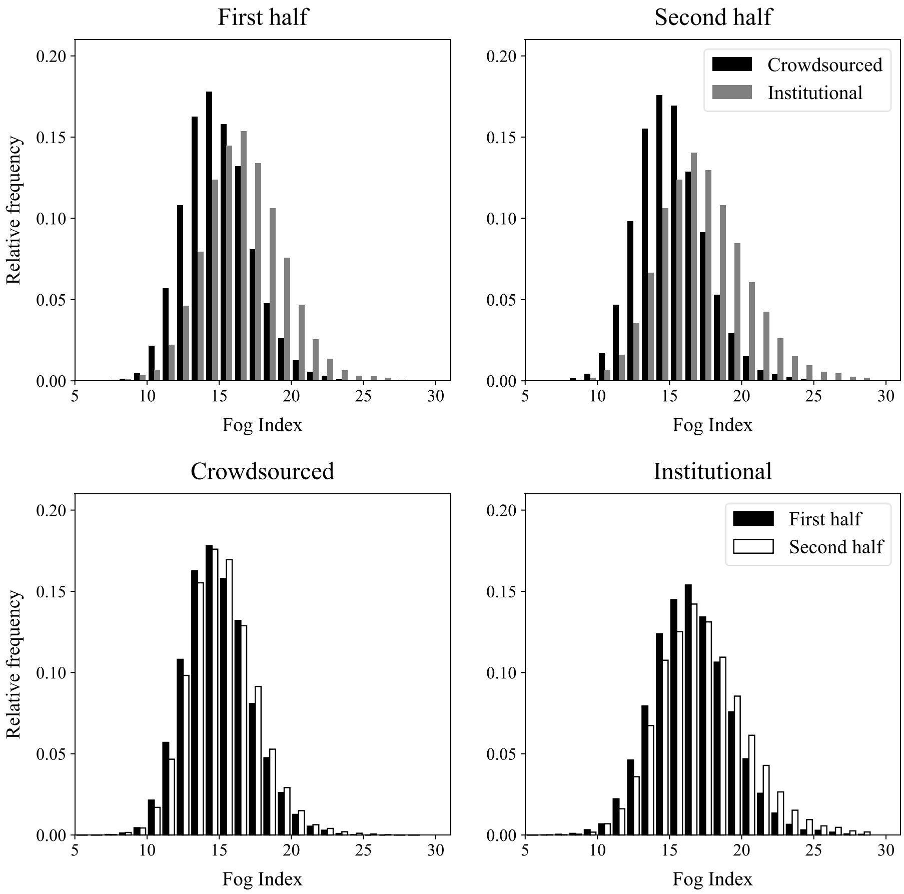

Figure 6. Readability score of analyst report bisection

But does this lead to the conclusion that PAs are better in writing "abstracts" compared to their NPA peers? Institutional analyst reports are overall more difficult to read. The PAs have therefore more room for improvement on the first pages. Not surprisingly, we see a stronger improvement in reports from PAs. Looking at the absolute level, on average the second half of the crowdsourced reports (mean score of 15.17) are written simpler than the first half of the institutional reports (mean score of 16.64).

Information Density

We discussed the readability of the two report types in detail but have not considered the content jet. Therefore, we calculate Shannon Entropy to evaluate the information density of the reports. To hold the text length constant, the analysis is based on the first 500 words of the report. We excluded reports with less than 500 words from our analysis. The remaining sample consists of 13,290 reports, with 3,311 crowdsourced and 9,979 institutional analyst reports. As shown in Figure 7, crowdsourced reports from NPAs have a higher Shannon Entropy compared to their professional counterparts. We observe a highly significant difference (p < 0.001) in mean Entropy of 0.0660. This implies that NPAs are able to provide even more information for a fixed text length in addition to easier readability. For this reason, we consider hypothesis **H4** to be confirmed.

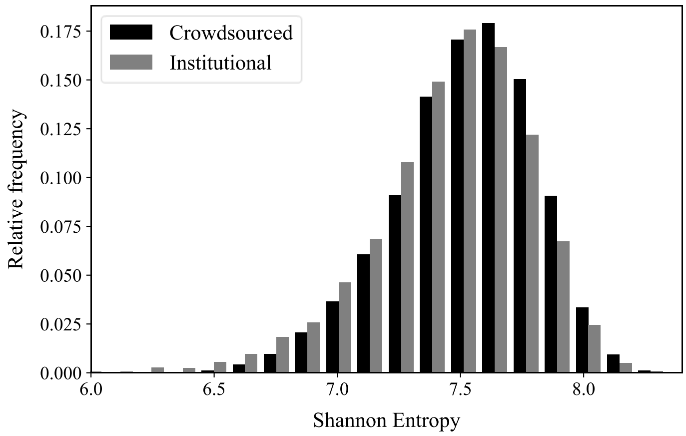

Figure 7. Shannon Entropy by document type

Discussion

Our study presents new insights into the readability of both institutional and crowdsourced analyst reports. Previous studies, e.g., De Franco et al. (2015) and Hsieh et al. (2016), find scores of 18.71 and 14.01 for the Fog Index of institutional analyst reports. Our result of 17.10 is within this range and therefore contributes to a better understanding of institutional analyst reports' readability. Our main findings reveal that NPAs provide similar content (**H1**) in an easier to understand format (**H2**) and even with a higher information density (**H4**) compared to their institutional peers. Also, the readability is much better in the first half of the documents than in the second half (**H3**). Thus, we contribute to a better understanding of NPAs. We can thereby also give an indication as to why NPAs have gained influence. This fits in with the fact that a previous study even found a market reaction on the publication of crowdsourced analyst reports (Campbell et al. 2019). We find that crowdsourced reports have a particularly high content similarity (measured by cosine similarity) with institutional reports that were published earlier. The similarity to institutional reports published later is rather low (Figure 3). In other words, on average, institutional analysts provide new information more quickly. This is in opposition to the paper of Drake et al. (2019), which states that NPAs process information faster in their reports. The extent to which the two types of analysts not only cover the same topics but also copy from each other could be considered in further analyses.

In Europe, the market for research conducted by PAs might become more competitive. This is because of the new MiFID II regulation, introduced by the European Commission in 2018. This regulation requires brokers to unbundle their execution services from their offered company research (Fang et al. 2019). Therefore, analyst reports provided by brokerage companies must be billed separately. Fang et al. (2019) found evidence that MiFID II pushed investment companies towards the insourcing of analyst research. Although MiFID II is a European regulation it has a global impact. Allen (2019) found higher competition in the US investment research marketplace since the MiFID II implementation. As the quality of their services becomes more important under these conditions, PAs might learn from NPAs how to delight their readers. The highly competitive environment on an online platform could be one reason why NPAs stick to the easier readability of their reports. The PAs might use the reports of NPAs as a reference point to improve their texts and be prepared for more competition within their market. Such improvements are somewhat limited in that PAs have to meet requirements on certain regulations regarding the contents of the reports and provide, for example, detailed explanations or research on non-core topics. The results of our study are limited due to the data sample. Since readability scores are sensitive to a proper word and sentence tokenization, the pre-processing has a major impact on the results. Even though we applied numerous plausibility checks and manually checked the data, we cannot guarantee that the tokenization is free of errors. To further enhance the robustness of the results, we calculated the readability measures by deleting stopwords. Although the level of these measures has shifted slightly, the differences between the two groups remain almost identical. Nevertheless, as already described, we regard it as appropriate not to remove stopwords when calculating readability measures. Furthermore, we restricted the analysis to the text of the documents. However, analyst reports (institutional and crowdsourced) also contain a substantial number of tables, charts, and figures. These elements also support a better text understanding. As the inclusion of these document elements would be very complex, the influence of these elements has not been considered. It seems possible that PAs use complicated language to generate a more sophisticated impression and thereby unintentionally foster obfuscation. However, the authors' intention is not within the scope of our analysis, as we want to focus on the information transfer efficiency.

Existing literature indicates stronger negative biased stock forecasts for PAs than for NPAs (Jame et al. 2016). We view this as a future research opportunity for an analytical approach on the textual level that is comparable to the one in our paper. First analyses with our dataset show that both groups of analysts have a comparably high subjectivity score, but the average sentiment score for institutional reports is five times higher than the score for crowdsourced reports. These preliminary results indicate more positive biased forecasts of PAs and stand in contrast to the findings of Jame et al. (2016).

Conclusion

The importance of sell-side analyst reports has declined in recent years (Drake et al. 2019). On the one hand, this can be explained by investment banks and fund companies increasingly turning to own research. On the other hand, this may be related to research being available free of charge on the Internet. The latter type of research is carried out by individuals who do not belong to a brokerage house but who have expertise in the area of finance and regarding a company or industry. Recent research investigates the increasing role of crowdsourced content and demonstrates that NPAs contribute to capital market efficiency (Campbell et al. 2019). Referring to the behavior pattern of obfuscation (Li 2008), the above-mentioned developments can be explained by information being presented differently. We analyze whether the texts of different kinds of analyst reports provide an explanatory contribution.

In a first step, we check for the analyst reports that they are covering the same content. Subsequently, we calculate readability scores for the texts. We show that crowdsourced analyst reports are much easier to read than institutional analyst reports. We further confirm this result by comparing individual report pairs that target the same company and are chronologically as close as possible. Also, we calculate the information density of the texts utilizing the concept of Entropy. We show that crowdsourced reports contain more information within the same text length compared to institutional reports. Given the above-mentioned results, we conclude that information provision differs between texts of institutional and crowdsourced analyst reports. Therefore, crowdsourced analyst reports are capable of reducing information processing costs. To convey information better, it might be advantageous for PAs to shorten their very long sentences (see Figure 2) and to present the contents as concisely as possible, if reasonably feasible.

It should be noted that our results depend to a certain degree on the pre-processing of the texts. However, we show in our study through various evaluations that the numbers of sentences, words, and complex words are within a reasonable range. Since our results are based on a limited dataset, expanding the dataset and considering industries could provide further insights. Also, the concept of Entropy should only be regarded as an indication of information density. Nevertheless, we consider the use of Entropy to be appropriate since our approach controls for the document length to compare report types, making the results more robust. Our analysis contributes to an improved understanding of the increasing success of crowdsourced research. We also confirm the existing research findings on the readability of institutional analyst reports. PAs might consider crowdsourced analyst reports as a reference point to make their reports more reader-friendly and maintaining information content simultaneously. With this paper, we elaborate on the difference between PAs and NPAs in how they present information to their readers. However, this is most likely not the only difference between the two groups of analysts. Therefore, further studies are necessary, which in particular highlight differences in content.

Appendix

Table 2. Dataset

Company	Crowdsourced		Institutional Reports	
	N Reports	N Authors	N Reports	N Authors
3M	129	75	320	31
AT&T	276	108	613	49
American Express	62	34	673	42
Apple	881	237	1,337	77
Boeing	268	73	703	42
Caterpillar	184	88	735	39
Chevron	252	72	474	42
Cisco Systems	191	84	809	53
Coca-Cola	185	99	611	41
DowDuPont	42	26	442	31
Exxon Mobil	360	130	445	41
General Electric	703	164	699	44
Goldman Sachs Group	61	36	482	39
Home Depot	131	67	520	42
Intel	166	78	872	65
IBM	340	149	473	52
JPMorgan	136	66	604	41
Johnson & Johnson	259	116	760	45
McDonald's	230	119	589	47
Merck	43	26	610	39
Microsoft	304	148	711	57
Nike	190	112	659	49
Pfizer	75	49	544	41
Procter & Gamble	168	85	418	41
Travelers Companies	40	17	394	34
United Technologies	38	23	438	34
UnitedHealth Group	21	17	575	38
Verizon	123	70	599	47
Visa	86	51	543	42
Walgreens Boots Alliance	43	32	507	32
Walmart	243	126	775	52
Walt Disney	245	132	484	53
Total	6,475	932*	19,418	164*
Mean	202.34	84.66	606.81	44.44

*unique authors within the entire sample

References

Allen, J. 2019. "US Payment for Research and MiFID II 2019 Survey," *CFA Institute*.

Annam, M., and Sajeev, G. P. 2016. "Entropy Based Informative Content Density Approach for Efficient Web Content Extraction," in: *Proceedings of the 2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, Jaipur, India, pp. 118–124.

Asay, H. S., Elliott, W. B., and Rennekamp, K. 2016. "Disclosure Readability and the Sensitivity of Investors' Valuation Judgments to Outside Information," *The Accounting Review* (92:4), pp. 1–25.

Asquith, P., Mikhail, M. B., and Au, A. S. 2005. "Information Content of Equity Analyst Reports," *Journal of Financial Economics* (75:2), pp. 245–282.

Aurnhammer, C., and Frank, S. L. 2019. "Evaluating Information-Theoretic Measures of Word Prediction in Naturalistic Sentence Reading," *Neuropsychologia* (134).

Brown, L. D., Call, A. C., Clement, M. B., and Sharp, N. Y. 2015. "Inside the 'Black Box' of Sell-Side Financial Analysts," *Journal of Accounting Research* (53:1), pp. 1–47.

Campbell, J. L., DeAngelis, M. D., and Moon, J. R. 2019. "Skin in the Game: Personal Stock Holdings and Investors' Response to Stock Analysis on Social Media," *Review of Accounting Studies* (24:3), pp. 731–779.

Chen, X., Cheng, Q., and Lo, K. 2010. "On the Relationship between Analyst Reports and Corporate Disclosures: Exploring the Roles of Information Discovery and Interpretation," *Journal of Accounting and Economics* (49:3), pp. 206–226.

Courtis, J. K. 1986. "An Investigation into Annual Report Readability and Corporate Risk-Return Relationships," *Accounting and Business Research* (16:64), pp. 285–294.

Courtis, J. K. 1995. "Readability of Annual Reports: Western Versus Asian Evidence," *Accounting, Auditing & Accountability Journal* (8:2), pp. 4–17.

De Franco, G., Hope, O. K., Vyas, D., and Zhou, Y. 2015. "Analyst Report Readability," *Contemporary Accounting Research* (32:1), pp. 76–104.

Dethlefs, N., Hastie, H., Rieser, V., and Lemon, O. 2012. "Optimising Incremental Dialogue Decisions Using Information Density for Interactive Systems," in: *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL),* Jeju Island, Korea: pp. 82–93.

Drake, M. S., Moon, J., Twedt, B. J., and Warren, J. 2019. "Are Social Media Analysts Disrupting the Information Content of Sell-Side Analysts' Reports?," *available at SSRN 3456801*.

Fang, B., Hope, O.-K., Huang, Z., and Moldovan, R. 2019. "The Effects of MiFID II on Sell-Side Analysts, Buy-Side Analysts, and Firms," *available at SSRN 3456801*.

Flesch, R., and Gould, A. J. 1949. *The Art of Readable Writing*. New York: Harper.

Frankel, R., Kothari, S., and Weber, J. 2006. "Determinants of the Informativeness of Analyst Research," *Journal of Accounting and Economics* (41:1-2), pp. 29–54.

Gunning, R. 1952. *The Technique of Clear Writing*. New York: McGraw-Hill.

Hsieh, C. C., Hui, K. W., and Zhang, Y. 2016. "Analyst Report Readability and Stock Returns," *Journal of Business Finance & Accounting* (43:1-2), pp. 98–130.

Huang, A. H., Lehavy, R., Zang, A. Y., and Zheng, R. 2017. "Analyst Information Discovery and Interpretation Roles: A Topic Modeling Approach," *Management Science* (64:6), pp. 1–23.

Jame, R., Johnston, R., Markov, S., and Wolfe, M. C. 2016. "The Value of Crowdsourced Earnings Forecasts," *Journal of Accounting Research* (54:4), pp. 1077–1110.

Jame, R., Markov, S., and Wolfe, M. 2017. "Does Crowdsourced Research Discipline Sell-Side Analysts?," in: *SMU Cox School of Business Research Paper* (18-6).

Kincaid, J. P., Fishburne Jr., R. P., Rogers, R. L., and Chissom, B. S. 1975. "Derivation of New Readability Formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel," *Institute for Simulation and Training, University of Central Florida*.

Lawrence, A. 2013. "Individual Investors and Financial Disclosure," *Journal of Accounting and Economics* (56:1), pp. 130–147.

Lee, J. 2019. "Analyst Jobs Vanish as a Perfect Storm Crashes into Research." Retrieved 02-25-2020, https://www.bloomberg.com/news/articles/2019-12-19/analyst-jobs-vanish-as-a-perfect-storm-hits-wall-street-research

Li, F. 2008. "Annual Report Readability, Current Earnings, and Earnings Persistence," *Journal of Accounting and Economics* (45:2-3), pp. 221–247.

Loughran, T., and McDonald, B. 2014. "Measuring Readability in Financial Disclosures," *Journal of Finance* (69:4), pp. 1643–1671.

Rennekamp, K. 2012. "Processing Fluency and Investors' Reactions to Disclosure Readability," *Journal of Accounting Research* (50:5), pp. 1319–1354.

Shannon, C. E. 1948. "A Mathematical Theory of Communication," *Bell System Technical Journal* (27:3), pp. 379–423.

Štajner, S., Evans, R., Orasan, C., and Mitkov, R. 2012. "What Can Readability Measures Really Tell Us About Text Complexity," in: *Proceedings of the the 2012 Workshop on Natural Language Processing for Improving Textual Accessibility (NLP4ITA),* Istanbul, Turkey, pp. 14–21.

Womack, K. L. 1996. "Do Brokerage Analysts' Recommendations Have Investment Value?," *Journal of Finance* (51:1), pp. 137–167.