Received: 21 September 2022 | Revised: 21 November 2023 | Accepted: 6 December 2023

DOI: 10.1111/ijau.12343

ORIGINAL ARTICLE

WILEY

Back to where it started?—Do expanded auditor's reports become sticky, generic and boilerplate over time?

Andreas Seebeck

School of Business, Social & Decision Sciences, Constructor University, Bremen, Germany

Correspondence
Andreas Seebeck, School of Business, Social & Decision Sciences, Constructor University Bremen, Campus Ring 1, 28759 Bremen, Germany.
Email: a.seebeck@jacobs-university.de

Funding information
This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors

[Correction added on 1 February 2024, after first online publication: Projekt DEAL funding statement has been added.]

Although auditing standards instruct auditors to avoid issuing standardised key audit matters (KAMs), standard setters warn and financial statement users worry about expanded auditor's reports becoming sticky, generic and boilerplate over time. Based on textual analysis of expanded auditor's reports from the UK, I provide longitudinal evidence of an increase in various textual similarity measures including year-over-year similarity (stickiness), within-peer group similarity (generality) and boilerplate language over time. Next, I identify a number of crucial drivers of convergence and find that disclosure similarity diminishes the positive capital market effect of precise KAM reporting documented in prior studies. Together, these findings suggest that audit firms have successively developed standard text modules for KAM reporting that undermine the aim of expanded auditor's reports to provide informative value. This paper extends the literature on the relationship between institutional environments and unintended effects of auditor disclosure regulation.

KEYWORDS
auditor reporting regulation, boilerplate language, critical audit matters (CAMs), expanded auditor's reports, key audit matters (KAMs), similarity

1 | INTRODUCTION

The binary 'pass-or-fail' model of auditor's reports was long criticised for containing mostly generic and boilerplate language and providing only limited insight into the audit of financial statements (e.g. Christensen et al., 2019; Gray et al., 2011). In response to this criticism, several standard setters, including the UK Financial Reporting Council (FRC), developed new reporting frameworks, expanding auditor's reports by client- and engagement-specific information (e.g. FRC, 2013a; IAASB, 2015; PCAOB, 2017). To this end, revised 'ISA 700 (UK & Ireland)' from June 2013, henceforth referred to as ISA 700, was one of the first standards that required auditors to report the risks of material misstatement (RMMs) that had the greatest effect on the audit. These RMMs, which later became the key audit matters (KAMs),[1] must be communicated '*in a way that enables them to be related directly to the specific circumstances of the audited entity and are not, therefore, generic or abstract matters expressed in standardized language*' (ISA 700.19B).[2] However, persisting doubts and concerns about the actual informativeness of KAMs are what motivates this study (Bédard et al., 2016).

Already prior to the implementation of ISA 700, scholars, regulators, users of financial statements, and auditors expressed concern about whether client- and engagement-specific information would be provided in KAMs (e.g., FRC, 2013b; Lennox et al., 2023; Norris, 2014). For instance, the second most raised concern in response to an FRC consultation paper on the revised ISA 700 standard was the expected use of boilerplate language in KAM reporting (FRC, 2013b). In line with this anxiety, post-implementation reviews conducted by the FRC revealed a significant amount of generic KAM disclosures and a lack of change in KAM reporting between periods in the initial

years (FRC, 2015, 2016c). However, comprehensive cross-sectional evidence on the similarity of KAM disclosure is missing, as the sample periods of most prior studies are rather short, and textual similarity of KAMs has not yet been investigated in detail.[3] Together, the concerns and findings described above call for more in depth research in this field.

This paper examines different textual similarity measures of expanded auditor's reports and KAM descriptions issued in the first 6 years after the implementation of ISA 700 to study the long-term development of expanded auditor disclosure similarity, its drivers and potential capital market consequences. It uses three different similarity measures that assess different types of textual similarity including year-over-year similarity (stickiness), within-peer group similarity (generality) and the use of boilerplate language.[4]

The first similarity measure, 'stickiness', refers to the similarity of reporting over time. In line with Dyer et al. (2017), 'stickiness' measures year-over-year language changes in disclosure, where fewer changes in language result in stickier disclosure. Hence, greater stickiness indicates a lack of period-specific reporting. In contrast, the 'within-peer group similarity' measures the similarity of a given KAM description with all other KAM descriptions that have (a) been reported in the same reporting period, (b) been reported by the same audit firm in the same reporting period, or (c) describe the same type of risk for the same reporting period. Thus, a high 'within-peer group similarity' score would indicate that auditors refrain from using a lot of client-specific information but rather use audit firm or risk-type generic disclosure. The use of 'boilerplate language', as the third similarity measure, denotes the use of standard language across a variety of KAM descriptions. The use of boilerplate language is often seen as a negative attribute because it does not provide investors with detailed information about a company's unique circumstances. In line with prior studies, I define boilerplate disclosure as '*standard disclosure that uses many words with little firm-specific or fiscal-period-specific content*' (Brown & Tucker, 2011, p. 312) and which is used pervasively among reports and consequently provides no client-specific information (e.g. Cazier et al., 2021; Hope et al., 2016; Lang & Stice-Lawrence, 2015).

All three measures have their strengths and weaknesses in identifying the similarity of reporting. They are all effective and commonly used for measuring textual similarity, but neither can be unequivocally declared as superior to the other. For instance, stickiness and within-peer group similarity, which are measured using the Cosine similarity approach,[5] are less affected by document length and robust to variations in word order. Hence, documents that use different word orders to express the same ideas will still be considered similar. In contrast, the boilerplate similarity measure is determined using the n-gram approach, which assesses textual similarity based on the sequential overlap of groups of n words between the two texts, therefore taking into account the order of the words.[6] Consequently, it is sensitive to slight changes in the text, such as a word order change or replacements of words and, thus, can be used to capture local syntactic similarity. The three approaches can be used together to provide a holistic measure of text similarity.

Gathering empirical evidence on how investors react to the similarity of KAM disclosure might help to further our understanding of precisely what makes KAM informative. For instance, if auditors stick to their previous reports or use similar disclosures across their audit clients rather than providing client- and engagement-specific disclosures, auditor's reports do not change appreciably from the previous year and do not differ substantially from other firms' auditor's reports. Thus, they might not serve their intended purpose. Similarly, if the auditors use boilerplate language, investors might lack detailed client- and engagement-specific information and thus find it more difficult to make informed decisions.

Therefore, the empirical strategy of this paper is twofold. First, I investigate the development of KAM disclosure similarity over time. The unique setting of the UK provides one of the longest available time-series samples of expanded auditor's reports including KAMs.[7] I find robust evidence that year-over-year similarity of expanded auditor's reports increases over time, showing that auditors increasingly stick to KAM disclosures from the previous year. Further, I find that all of the aforementioned measures of 'within-peer group similarity' as well as the use of 'boilerplate language' rise over time. Recurring disclosure does not necessarily need to be bad, as some potentially important information on persisting matters must sensibly be carried forward from year to year. Its removal would lead to a loss of reporting quality (Miller, 2017). Likewise, within-peer group similarity of KAM disclosure can be driven by similar circumstances and risk environments faced by the auditors during the audit and thus be suitable. However, ex ante there is little reason why disclosure similarity should increase over time. The documented growth in similarity suggests that auditors increasingly use standard text modules when reporting KAMs, resulting in less client- and reporting period-specific disclosure.

In the second step, I investigate the capital market effects of the similarity of KAM disclosure. Early evidence from experimental studies using laboratory settings indicates that expanded auditor's reports can provide useful information to financial statement users (e.g. Brasel et al., 2016; Christensen et al., 2014; Gimbar et al., 2016). The prominent position of the auditor's report in financial statement documents the higher source credibility of the auditor's reports over management disclosure, as well as the limited attention of investors can explain why KAMs can be informative and useful despite investors' prior access to the information via alternative channels.[8] However, the results of existing empirical archival studies that test for the informativeness and usefulness of KAMs are heterogenous (e.g. Gutierrez et al., 2018; Lennox et al., 2023; Porumb et al., 2021; Reid et al., 2019; Seebeck & Kaya, 2022). While Gutierrez et al. (2018) and Lennox et al. (2023) find no association between KAM disclosure and capital market reactions, Seebeck and Kaya (2022) find that more precise descriptions of KAMs are significantly and positively associated with cumulative abnormal returns and abnormal trading volume around the filing dates. Moreover, Reid et al. (2019) find that KAM disclosures are associated with improved financial reporting quality, and Porumb et al. (2021) find that they contain relevant information for loan contracting resulting in reduced loan spread and longer maturity for loan facilities.

This study fails to provide evidence of a direct relationship between KAM disclosure similarity and capital market effects. However, building on the findings of Seebeck and Kaya (2022), who provide evidence that more precise descriptions of KAMs are significantly and positively associated with capital market reactions, I find that disclosure similarity moderates this effect. The results suggest that investors value precise information in KAMs (i.e. names of persons, locations and organisations, as well as quantitative information such as percentages, money values, and times and dates) less if disclosure similarity is higher.

Overall, my results point to the conclusion that expanded auditor's reports converge over time, driven by institutional forces as well as auditor and client characteristics. Thus, they validate critics' lasting concerns surrounding a development towards standardised KAM reporting. Further, they are timely and relevant in light of the ongoing development of auditing standards worldwide. Notwithstanding the fact that in some circumstances standardisation in financial disclosure is useful and allows for easier benchmarking (e.g. Blankespoor et al., 2020; Kaya & Koch, 2015), homogeneity in KAM reporting bears the risk of auditor's reports becoming potentially misleading in their portrayal of the current year's audit and material risks. The associated negative consequences are reflected in the finding that higher similarity in KAM descriptions diminishes the positive capital market effect from precise disclosure as reported by Seebeck and Kaya (2022).

This study makes a few important contributions. First, existing research on the implications of expanded auditor reporting has predominantly focussed on the years immediately before and after the implementation of the new auditing standards. Many studies examine adoption effects by using within-firm, pre- to post-adoption period comparisons. By exploring the long-term development of expanded auditor reporting, I add to the existing research studying the circumstances under which users value the inclusion of KAMs (e.g. Moroney et al., 2021). Notably, I find that KAM disclosure similarity moderates the positive capital market effect of precise KAM disclosure.

Second, I extend the literature on unintended consequences of disclosure regulation (e.g. Bushee & Leuz, 2005; Kelton & Montague, 2018; Linck et al., 2009). Specifically, I show that auditors increasingly rely on standardised text modules when reporting KAMs. This standardisation undermines the goal of reporting audit matters for individual clients and distinct reporting periods. Consequently, the essence of my research provides actionable insights for regulators. By applying textual analyses to expanded auditor's reports, they can identify instances of sticky, generic, and boilerplate disclosures, and thereby equip themselves to take several specific actions to counteract these issues. For instance, they can conduct regular training sessions and workshops for auditors to reinforce the importance of entity- and reporting period-specific reporting in KAMs, implement review and feedback mechanisms such as peer reviews, and develop reporting standards and guidelines that promote useful KAM disclosure and counter disclosure overload by limiting non-client-specific legalistic and technical language.

Third, I add to the auditor reporting literature that uses advanced text mining techniques to study the effects of KAM reporting. While some pioneering auditing studies proxy for informative value of expanded auditor's reports by using textual characteristics (e.g. Seebeck & Kaya, 2022; Smith, 2023), a recent study by Burke et al. (2023) shows that textual characteristics of the management disclosure such as its length, tone and stickiness are driven by KAM reporting, suggesting that KAMs have an effect on management disclosure. Using different similarity measures, I test and find that the concerns regarding an increase in standardised language in KAM reporting over time are indeed valid, contrary to standard setters' intentions and despite the adverse consequences this may have on capital markets.

2 | REGULATORY BACKGROUND AND PRIOR RESEARCH

2.1 | Regulatory background

The UK regulator, the FRC, made significant changes to corporate and auditor reporting requirements in 2013. In this endeavor, the FRC was one of the first regulators to release expanded auditor reporting standards, that is, ISA 700 (UK & Ireland) 'The Independent Auditor's Report on Financial Statements' (Revised June 2013), effective for audits of financial statements for periods commencing on or after 1 October 2012 (ISA 700.6). ISA 700 sets three new high-level requirements for the auditor's report of companies with a premium listing of equity shares on the London Stock Exchange that aim to provide users of financial statements with more transparent information. Most importantly, it requires auditors to identify and disclose the assessed 'RMMs', which had the greatest effect on the audit engagement (ISA 700.19A). In addition, information on the scope of the audit and the concept of materiality, including how the auditor applied materiality in planning and performing the audit, as well as the materiality level, need to be disclosed.

In June 2016, the FRC again revised its auditor reporting standards to align with the IAASB's recently approved respective standards ISA 700 and ISA 701 (FRC, 2016a, 2016b; IAASB, 2015). Henceforth, there was an own standard for KAM reporting in the UK (i.e. ISA 701) and the term 'risks of material misstatement' (RMMs) was replaced with 'key audit matters' (KAMs), defined as '*those matters that, in the auditor's professional judgement, were of most significance in the audit of the financial statements of the current period*' (ISA 701.8). The alignment with the international standards does not bring significant changes as the definitions of RMMs and KAMs are broadly equivalent, which is also explicitly stated by the FRC (2016a). However, the revised standards ISA 700 and ISA 701 (June 2016) provide more guidance for auditors on how to determine KAMs, which can impact the wording of KAM descriptions and thus textual similarity.[9]

The most recent version of the standards is dated to May 2022 (FRC, 2022), which shows that the FRC regularly updates its standards and thus should be interested in the findings of empirical studies, including studies on the long-term consequences of existing auditor

reporting standards. Building on the long-term experiences in the UK can be a useful guide also for other standard setters and directly speaks to their review initiatives (e.g. FRC, 2015, 2016c; Irish Auditing and Accounting Supervisory Authority [IAASA], 2020).[10] For instance, conducting similarity analyses on expanded auditor's reports can help standard setters identify areas where sticky, generic and boilerplate disclosure may be occurring.

2.2 | Prior research

2.2.1 | Research on textual similarity in financial disclosure literature

Textual analysis is becoming increasingly popular in empirical accounting research (Loughran & McDonald, 2016, 2020; Seebeck & Vetter, 2022). Besides 'length', 'readability', 'specificity' and 'sentiment', 'similarity' is among the most frequently examined textual attributes of financial reports (e.g. Dyer et al., 2017; Lang & Stice-Lawrence, 2015). Existing studies examine the similarity of various types of financial reports such as IPO prospectuses (e.g. Hanley & Hoberg, 2010), analyst reports (e.g. Huang et al., 2014), and specific sections and elements of annual reports (e.g. Brown & Tucker, 2011; Dyer et al., 2017; Lang & Stice-Lawrence, 2015). For instance, Brown and Tucker (2011) provide evidence for a decline in MD&A year-over-year modification scores (i.e. higher year-over-year similarity) in a sample period from 1997 to 2006, accompanied by longer disclosures and lower price reactions. Their findings suggest that MD&A sections become more repetitive and provide less informative value over the years. In line with this notion, Dyer et al. (2017) show that there is considerable stickiness in the disclosure of 10-K filings of US firms. Moreover, Lang and Stice-Lawrence (2015) provide evidence of increasing boilerplate reporting in the annual reports of both US and non-US firms.

Overall, prior research suggests that *firms* providing sticky, generic and boilerplate disclosures experience negative capital market consequences, such as higher cost of capital, greater stock price volatility and weaker market responses, as well as a decline in the ability of analysts to assess fundamental risk (e.g. Campbell et al., 2014; Hope et al., 2016; Kravet & Muslu, 2013). Consequently, regulators repeatedly point out that management should avoid boilerplate disclosures in favour of disclosures that are specific to the firm (e.g. SEC, 1998, 2005, 2016). Similarly, the FRC instructs auditors to avoid generic and standardised language in KAM disclosure (ISA 701. A47). However, ex ante, the capital market effects of standardised disclosure in KAMs remain unclear, as does the question of what drives the development of textual similarity over time.

2.2.2 | Research on textual characteristics of KAMs

A wide range of recent studies investigate the informativeness and usefulness of KAM reporting for financial statement users.[11] Most experimental studies suggest that the expanded auditor's report can be useful by providing incremental information (Christensen et al., 2014), adding credibility to existing information (Elliott et al., 2020), directing attention to matters of high risk (Sirois et al., 2018) and providing a forewarning for potential audit issues (Asbahr & Ruhnke, 2019). However, the findings of empirical archival studies are more heterogeneous with respect to the research questions of whether KAM disclosures provide informative value (Gutierrez et al., 2018; Lennox et al., 2023; Reid et al., 2019; Seebeck & Kaya, 2022). Examining textual characteristics of expanded auditor's reports including their textual similarity seems a promising way to better understand the reasons for these contrary findings.

There are a few studies using textual analyses to investigate expanded auditor's reports. For instance, Pinto et al. (2020) find that the disclosure of KAMs that are based on accounting standards with higher rule-based characteristics decreases the readability of auditor's reports. Further studies find that text properties such as length, readability and tone of auditor's reports change in the first year of ISA 700 reporting (e.g. Burke et al., 2023; Smith, 2023). In a recent study, Seebeck and Kaya (2022) provide pioneering evidence that many textual characteristics of expanded auditor's reports including readability, evaluative content, visual aids and specificity further increase in the second and third year after implementation. However, advancements in the third year are found to be only minor. Therefore, the authors consider that the improvement of auditor reporting may stagnate over time as disclosure converges. This thesis, however, remains unexplored in their study and a gap in literature.

Prior studies also find null evidence that most of the textual characteristics of expanded auditor's reports and KAM reporting are associated with capital market reactions (e.g. Lennox et al., 2023; Seebeck & Kaya, 2022; Smith, 2023). So far, only the specificity of KAM descriptions was found to be significantly and positively associated with equity market reactions (i.e. higher absolute cumulative abnormal returns and higher abnormal trading volume around the financial statement release dates). While disclosure can be precise, as according to the specificity measures used in prior disclosure studies, it can simultaneously be similar to that of the previous year or to the disclosure of a peer firm. Thus, understanding the capital market effects of textual similarity of KAM disclosure as well as the interplay between textual similarity and specificity is of high relevance for standard setters and financial statement users.

To the best of my knowledge, there are so far only few studies that deal with textual similarity of expanded auditor's reports. First, Burke et al. (2023) provide pioneering evidence for the US setting, showing how the similarity of auditor's report language significantly decreased following the implementation of expanded reporting requirements. However, they find that the similarity of KAM disclosure significantly varies between auditor's reports issued in the first year after implementation. Second, Zeng et al. (2021) find for the Chinese setting that textual similarity differs between different KAM components, that is, auditors use less similar language in KAM descriptions than in the reporting of their reactions. Moreover, their findings suggest that textual similarity of KAMs is negatively

related to clients' earnings quality and the propensity of an unmodified opinion. Third, Carlé et al. (2023) find that KAM disclosure similarity is negatively correlated to an audit firm change. Finally, in a recent working paper, Deneuve et al. (2023) provide initial evidence that temporal as well as cross-sectional dissimilarities in KAMs can be informative for investors. However, overall, the drivers and consequences of KAM disclosure similarity remain largely unexplored.

3 | HYPOTHESES DEVELOPMENT

3.1 | Does expanded auditor reporting become sticky, generic and boilerplate over time?

I draw on institutional theory to motivate my first hypothesis. In line with this theory, auditors can achieve legitimacy through conformance with societal values (Meyer & Rowan, 1977; Parsons, 1956, 1960). DiMaggio and Powell (1983) describe the resulting process of homogenisation as isomorphism. To this end, there are three drivers: coercive isomorphism, mimetic isomorphism and normative isomorphism. I expect all three to affect the development of KAMs' textual similarity over time.

According to coercive isomorphism, organisations (i.e. audit firms) and individuals (i.e. audit engagement partners) adopt similar practices over time to enhance their legitimacy (DiMaggio & Powell, 1983). Coercive isomorphism is derived from the pressure placed by outside parties on reliant organisations and can be driven by legal and regulatory requirements. These requirements can create sticky auditor's reports as there are only little changes in the standards over time. Moreover, they create a standardised framework for reporting, which can affect the extent to which audit firms tailor their disclosures to their clients' specific circumstances, rather than relying on generic or boilerplate language. For example, if guidance explicitly requires auditors to provide detailed information on specific dimensions of KAMs, auditors may be more likely to provide unique and specific disclosures that reflect the individual circumstances of the audit. This could result in less similarity of disclosure across reports, as auditors respond to the standards in different ways based on their clients' specific risks and uncertainties. On the other hand, if guidance in the standards is too prescriptive or rigid, it could lead to a greater similarity of disclosure across auditor's reports, as auditors may follow a more standardised approach to reporting. This could result in less informative or less relevant disclosures, as audit firms may align their wording more closely with the formulations in the standards, thus increasing the use of boilerplate phrases, at the expense of firm-specific details and resulting in a convergence over time.

Mimetic isomorphism occurs through pressure on an organisation to imitate other organisations (DiMaggio & Powell, 1983). It holds that organisations and individuals tend to imitate each other's behaviour when facing uncertainty. This process of homogenisation takes place over time. In the case of KAM reporting, auditors may observe the reporting practices of their peers, particularly those who have high reputations or are considered experts in their industries and adopt similar practices. This convergence can be driven by the desire to reduce uncertainty in a complex and ambiguous environment or to enhance their professional reputation by aligning with industry leaders. Over time, this imitation process leads to increased similarity in auditor reporting. The desire to reduce uncertainty can also lead to auditors sticking to their own reporting from previous years. If a previous auditor's report was well received by the market, this might incentivise auditors to keep their KAMs consistent. On the opposite, changing KAMs from year to year could be perceived as a change in a firm's risk profile or indicate inconsistency in the auditor's approach, which is particularly unintended in the face of ambiguity. In line with this notion, survey evidence indicates that many auditors indeed prefer standardised KAMs (Pelzer, 2021). This view is also widely reflected in the concept release comments of expanded auditor's reports in the United States, in which many audit firms initially expressed support for a more standardised implementation of KAM disclosures (Reilly, 2014).

Normative isomorphism stems from the influence of professional networks, associations and educational institutions that shape the values, norms and expectations within a particular field (DiMaggio & Powell, 1983). In the context of KAM reporting, normative isomorphism can be driven by professional accounting organisations and educational institutions that establish best practices and guidelines for the auditing profession. These institutions play a significant role in shaping the behavior and practices of auditors, promoting a shared understanding of what constitutes appropriate reporting (i.e. professional consensus). As a result, auditor reporting tends to converge over time as auditors adhere to the best practices and guidelines established by these influential institutions. Because of the fact that guidelines typically do not change frequently over time, they can also lead to greater reporting stickiness. In addition, ongoing professional trainings can reinforce existing practices, leading to more similar reporting.

Combined, institutional factors provide several reasons for an increase in KAM disclosure similarity over time. However, sticky, generic and boilerplate reporting was not the intended consequence of revised ISA 700, and if auditors act in line with this notion, there may not be a convergence over time. Moreover, firms operate in changing environments and so management disclosure is typically changing over time. Consequently, audit conditions and environments, and therefore KAM disclosure, are also expected to alter. Hence, I present my first hypothesis, separately for the three similarity dimensions, in the null form:

HYPOTHESIS 1a. Expanded auditor's reports' year-over-year similarity (stickiness) does not increase over time.

HYPOTHESIS 1b. Within-peer group similarity of KAM disclosure does not increase over time.

HYPOTHESIS 1c. The use of boilerplate language in KAM disclosure does not increase over time.

3.2 | Does similarity of KAM disclosure have capital market effects?

Initial empirical archival studies fail to provide evidence for significant capital market reactions of the length of the expanded auditor's report and the number of unique KAMs disclosed (e.g. Gutierrez et al., 2018; Lennox et al., 2023). Recent studies, however, find that more advanced measures of KAM disclosure are associated with reduced loan spread and longer maturity for loan facilities (Porumb et al., 2021), as well as equity market reactions (Seebeck & Kaya, 2022). It is not clear, ex ante, whether more similar KAMs would also result in capital market effects.

On the one hand, the disclosure of similar KAMs can be seen as a positive signal from the auditor to the investors, helping to reduce information asymmetry between the two parties, as it can facilitate the comparison of firm risks over time and across an industry or sector. This could help investors better assess the relative risk profiles of different reporting periods and firms, leading to a more informed decision-making process and potentially more positive market reactions. In addition, more similar KAM disclosures can reduce uncertainty among investors. For instance, greater year-over-year similarity can indicate a stable risk profile of the firm. Moreover, greater within-peer group similarity and the use of boilerplate language may indicate that auditors have identified common risks that are prevalent across companies or industries. If these trends or risks are perceived as manageable or within the normal range of expectations, investors might view similar KAMs as a positive signal, indicating that firms are addressing long-term and industry-wide challenges effectively. Finally, when KAM disclosures are similar over time and across different auditor's reports, it could be seen as a signal that auditors are applying consistent standards and methodologies in their work. This may strengthen the credibility of auditors, making investors more confident in the quality of the financial statements and the audit. This increased trust may lead to positive market reactions.

On the other hand, higher KAM disclosure similarity results in a lack of differentiation. Consequently, investors may find it more difficult to track the development of certain risks and to differentiate between high- and low-risk firms, which could lead to increased uncertainty and skepticism among investors, resulting in negative market reactions as they reassess their investment positions. Moreover, similar KAM disclosures among peers and the use of boilerplate language might raise concerns about the possibility of contagion effects within an industry or sector. If one firm faces significant issues related to a disclosed KAM, investors may worry that other companies with similar KAM disclosures may be similarly affected. This fear could result in negative market reactions as investors adjust their portfolios to account for the perceived increased risk.

Overall, there are several reasons for and against the effect of KAM similarity on capital markets. The second hypothesis stated in the null form is:

> **HYPOTHESIS 2.** The similarity of KAM disclosure is not associated with capital market reactions.

3.3 | Does similarity diminish capital market effects of specific KAM disclosure?

Prior studies find that higher precision of KAM disclosure results in greater assimilation of the information, thus facilitating the incorporation of information into stock prices and resulting in stronger market reactions (Seebeck & Kaya, 2022). These findings are in line with the difficulty faced by investors extracting useful information from expanded auditor's reports, when the reporting is characterised by unspecific phrases (Lennox et al., 2023). They are also highly relevant for this study because prior literature directly relates specificity to similarity. For instance, Brown and Tucker (2011, p. 312) define boilerplate disclosure as '*standard disclosure that uses many words with little firm-specific or fiscal-period-specific content*', and Kravet and Muslu (2013) find that management disclosures that deviate from those of the industry peers are more likely to be firm specific. However, precise information, as measured by the specificity variables typically used in prior disclosure studies, does not necessarily mean that the reporting is dissimilar to that of other firms or that auditors avoid repetition from previous years.

For example, the following risk description is considered to be precise as according to standard specificity measures, because details about the time, names, locations, money values and percentages are provided: '*The outbreak of COVID-19 in the UK continued to impact the Group during the year ended 31 March 2021. The coronavirus (COVID-19) lockdown restrictions in 2020 saw a record £126.9 billion increase in bank deposits from households. These "forced savings" led to a rise in household savings of a record 23.9% in Quarter 2 (Apr to June) 2020*'. However, the disclosure is not firm-specific and, thus, can easily be used for various audit clients. Vice versa, dissimilarity does not necessarily mean that a KAM description provides precise and actually useful information.

Hence, the third hypothesis focuses on the interplay of similarity and specificity. I assume that high textual similarity is detrimental to KAM disclosure's ability to provide useful information and, thus, diminishes the positive capital market effects from specific disclosure as reported in prior literature. The third hypothesis is:

> **HYPOTHESIS 3.** Similarity diminishes the positive capital market effects from specific KAM disclosure.

4 | SAMPLE AND RESEARCH DESIGN

4.1 | Sample selection

Table 1A presents the sample selection process. I start by identifying all the firms listed on the FTSE 350 index as of 30 June 2016. I hand-collect the annual reports for the financial years 2012 to 2018 as provided on firms' corporate websites and by Thompson Reuters Datastream.[12] I then manually extract the auditor's reports including KAM sections along with additional auditor data.[13] Next, I convert all PDF documents to text files using PyPDF2 and pdfrw, which remove

TABLE 1 Sample selection and distribution.

Panel A: sample selection process	
Total available auditor's reports of FTSE 350 firms (as of 31 August 2016) for the observation period covering 1 year before and 6 years after the implementation of ISA 700	**2421**
Less auditor's reports for which no year-over-year similarity score could be calculated (e.g. missing comparison objects and technical issues)	*(444)*
Total available auditor's reports to calculate year-over-year similarity scores	1977
Less auditor's reports for financial services industry	*(690)*
Less pre-ISA 700 auditor's reports	*(180)*
Less observations with missing data for within-peer group similarity	*(24)*
Less observations with missing data for control variables	*(38)*
Final sample of extended auditor's reports	**1045**

Panel B: sample distribution by reporting period.			
Post-ISA 700 period	**Freq.**	**Percent**	**Cum.**
1	157	15.02	15.02
2	206	19.71	34.74
3	208	19.90	54.64
4	197	18.85	73.49
5	179	17.13	90.62
6	98	9.38	100.00
Total	**1045**	**100.00**	

Panel C: sample distribution by audit firm.			
Audit firm	**Freq.**	**Percent**	**Cum.**
BDO	13	1.24	1.24
Deloitte	277	26.51	27.75
EY	151	14.45	42.20
Grant Thornton	6	0.57	42.77
KPMG	296	28.33	71.10
PWC	302	28.90	100.00
Total	**1045**	**100.00**	

Notes: This table presents the sample selection procedure (Panel A) and summarises the sample distribution by post-ISA 700 reporting period (Panel B) and by audit firm (Panel C).

the majority of non-textual elements. This step allows for computer aided textual similarity analysis using Python's NLTK Package and RapidMiner Studio. I further conduct a thorough preprocessing of the extracted data including the removal of remaining formatting artifacts, tokenisation and conversion to lower case letters.

I eliminate all firms not covered by ISA 700.[14] In addition, I exclude investment trusts and firms with no available similarity data. Given that the year-over-year similarity requires two consecutive yearly observations for an audit client, I also eliminate auditor's reports for which no year-over-year similarity score can be calculated because of missing comparison objects and technical issues. Furthermore, I require firm-year observations to have full financial data available for all control variables. This results in a final sample of 1045 firm-year observations for non-financial services firms in the post-ISA 700 period, which is significantly larger than that of most existing UK-based studies on expanded auditor reporting. Tables 1B and 1C show the sample distribution among reporting periods and auditors, respectively.

4.2 | Textual similarity measures

Based on prior disclosure literature (e.g. Brown & Knechel, 2016; Brown & Tucker, 2011; Lang & Stice-Lawrence, 2015), this study uses three measures to evaluate textual similarity of expanded auditor's reports and KAM disclosure: (1) year-over-year similarity (stickiness), (2) within-peer group similarity of KAM disclosure (generality) and (3) the number of boilerplate phrases in KAM disclosure. All three measures are effective and commonly used for measuring textual similarity. Together, they provide a holistic view on the similarity of expanded auditor's reports.

4.2.1 | Year-over-year similarity (stickiness) of auditor's reports

In order to calculate the disclosure stickiness, I compare the reporting in a given year with the reporting of the previous year (e.g. Shell plc 2016 vs Shell plc 2015). Because of the lack of KAM disclosures in pre-ISA 700 periods, the stickiness of the *KAM sections* could only be calculated from the second post-ISA 700 year onward. This would lead to inconsistencies in the sample throughout the paper and significantly reduce the sample size for tests on *STICKINESS*. Moreover, it would exclude the particularly interesting first year of KAM disclosure, which serves to provide a holistic view of how disclosures have evolved over time, providing a backdrop against which changes in the post-ISA 700 period can be contextualised. Using the full auditor's report as corpus provides a workaround to measure the year-over-year similarity (*STICKINESS*) also for the first year after implementation of ISA 700, it compares an entire (expanded) auditor's report for a given client with its (expanded) auditor's report in the previous year.[15] A high score would indicate that auditors refrain from using a lot of reporting period-specific information but rather stick to the previous year's disclosure. *STICKINESS* is measured using the Cosine similarity approach as described in Appendix C. Figure 1 shows the development of *STICKINESS* over time for the year preceeding and the 6 years following the implementation of ISA 700.

4.2.2 | Within-peer group similarity of KAM disclosure

Following Brown and Knechel (2016), I use an extension of Brown and Tucker (2011) to examine textual similarity of KAM disclosure reported for different clients in the same reporting period. More precisely, I measure the similarity of a given KAM description to a peer

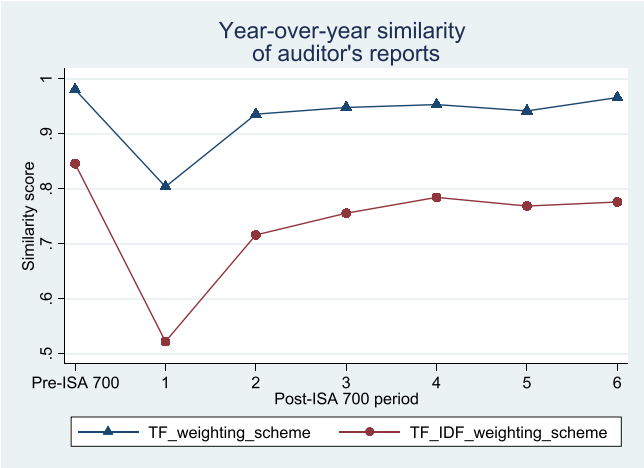

FIGURE 1 Year-over-year similarity of auditors reports: Figure 1 shows the development of year-over-year similarity (*STICKINESS*) over time for the year preceding and the 6 years following the implementation of ISA 700. *STICKINESS* was 84.5% (98.1%) when using the TF-IDF (TF) weighting scheme in the year before implementation. As expected, it sharply declined in the first year after the implementation of ISA 700 to 52.2% (80.4%) and steadily increased again in subsequent years. [Colour figure can be viewed at wileyonlinelibrary.com]

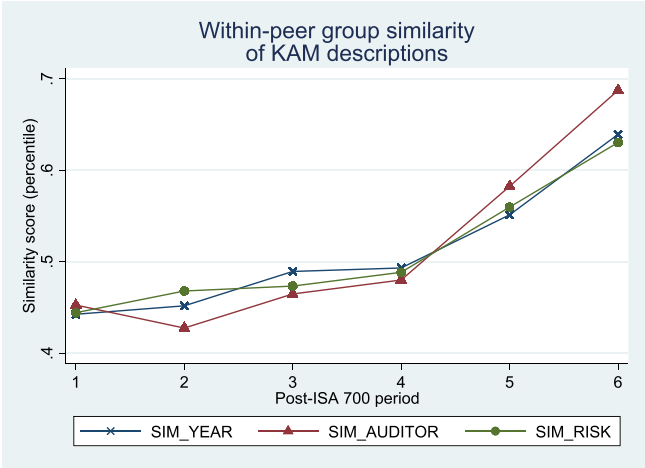

FIGURE 2 Within-peer group similarity of KAM descriptions: Figure 2 shows the development of within-peer group similarity of KAM descriptions (i.e. all KAMs that [a] have been reported in the same period (SIM_YEAR), [b] have been reported by the same audit firm in the same reporting period (SIM_AUDITOR) and [c] describe the same type of risk for the same reporting period (SIM_RISK)). All within-peer group similarity measures show a significant increase over time. [Colour figure can be viewed at wileyonlinelibrary.com]

group of all other KAM descriptions that have (a) been reported in the same period (*SIM_YEAR*), (b) been reported by the same audit firm in the same reporting period (*SIM_AUDITOR*), or (c) describe the same type of risk for the same reporting period (*SIM_RISK*).[16] While SIM_-YEAR can measure non-topic related standard wording and general auditor reporting trends, which can be for instance driven by explanations of the general KAM identification strategies required by the standards, SIM_AUDITOR can reflect changes in internal audit methodologies, guidelines and standard wording templates.

For each specification, I compute the mean similarity on auditor's report level by averaging the results for each KAM description contained in the report.[17] A high within-peer group similarity would indicate that auditors refrain from using a lot of client-specific information but rather use audit firm or risk-type generic disclosure. Technically, in the main analysis and in the robustness checks, I use the same approach as for the stickiness measure described in Appendix C. Figure 2 shows the development of within-peer group similarity of KAM descriptions over time.

4.2.3 | Boilerplate language in KAM disclosure

Finally, following Lang and Stice-Lawrence (2015), I examine the number of boilerplate phrases used in KAM disclosure. Boilerplate language consists of standard wording, which is used pervasively among reports and for this reason provides no client-specific information (e.g. Cazier et al., 2021; Hope et al., 2016; Lang & Stice-Lawrence, 2015). A greater use of boilerplate phrases in KAM disclosure would indicate that auditors refrain from providing a lot of specific information but rather use standardised disclosure.

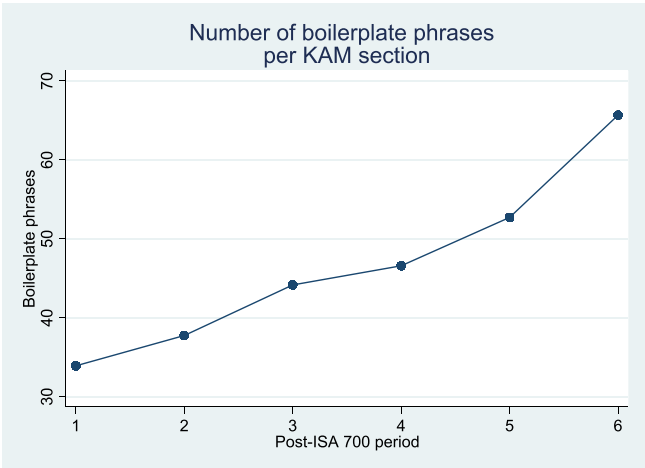

FIGURE 3 Number of boilerplate phrases per KAM section: Figure 3 shows the development of boilerplate phrases used in KAM sections (i.e. four-grams that are used frequently in KAM disclosure and for this reason provide no firm-specific information). [Colour figure can be viewed at wileyonlinelibrary.com]

In a two-step approach based on Lang and Stice-Lawrence (2015), I first identify sequences of four words—so-called 'four-grams'—which are found to be included regularly in the KAM sections of the sample firms (e.g. 'essential component [of] financial statements', 'significant amount [of] audit resources', and 'high risk [of] material misstatement').[18] Next, I count the number of boilerplate phrases for each KAM section. Figure 3 shows the development of boilerplate phrases used in KAM-sections over time. As this measure considers the order of words (i.e. the same four words in different

orders would be considered different four-grams), it detects similarity at a different level than the Cosine similarity score.[19]

4.3 | Research methodology

4.3.1 | Convergence of expanded auditor reporting over time (Hypotheses 1a-1c)

In order to examine whether there is an increase in textual similarity and boilerplate language over time, I adopt the OLS regression model of Seebeck and Kaya (2022) as presented in Equation (1).[20]

$$\begin{aligned} SIMILARITY_{i,t} = {} & \beta_0 + \beta_1\, TIMESREP_{i,t} + \beta_2\, SIZE_{i,t} + \beta_3\, ROA_{i,t} + \beta_4\, LEV_{i,t} \\ & + \beta_5\, MB_{i,t} + \beta_6\, MOWN_{i,t} + \beta_7\, LOSS_{i,t} + \beta_8\, CFO_{i,t} \\ & + \beta_9\, USLIST_{i,t} + \beta_{10}\, DELAY_{i,t} + \beta_{11}\, BIGN_{i,t} \\ & + \beta_{12}\, CHAUD_{i,t} + \beta_{13}\, EXPERT_{i,t} + \beta_{14}\, BUSY_{i,t} \\ & + Industry\ FE + \varepsilon_{i,t} \end{aligned} \tag{Eq1}$$

The dependent variable is the textual similarity of auditor's reports and KAM disclosure (*SIMILARITY*), measured by (1) year-over-year similarity of auditor's reports (*STICKINESS*), (2) three different within-peer group similarity measures of KAM descriptions (*SIM_YEAR*, *SIM_AUDITOR* and *SIM_RISK*), as well as an aggregate of these three measures (*SIM_PEER*), (3) the use of boilerplate language in KAM sections (*BOILERPLATE*) and (4) an aggregate similarity measure of 1 to 3 (*SIM_ALL*).[21] To construct the aggregate measures, I compute percentile ranks for each variable and then average them.[22] I also use these percentile rankings in place of interval measures in the main analyses, as they do not require a monotonic association with the independent variables and are easier to interpret (Loughran & McDonald, 2014). This research design choice does neither inflate the extent of variation between similarity scores nor distort the results.[23]

TIMESREP is a categorical variable that counts the number of times an expanded auditor's report has been issued under the new ISA 700 (701) regulation for a given audit client. It takes on the positive integer values of 1 to 6.

I control for various factors identified by Seebeck and Kaya (2022) that potentially affect the linguistic characteristics of auditor's reports.[24] These include client characteristics such as firm size (*SIZE*), profitability (*ROA*), leverage (*LEV*), market-to-book-ratio (*MB*), the occurrence of a loss (*LOSS*), change in operating cash flow compared to prior period (*CFO*), concentration of ownership of outstanding shares (*MOWN*), and cross-listings in the US (*USLIST*), as well as audit and auditor's characteristics such as the number of calendar days between a firm's fiscal year-end and auditor's report date (*DELAY*), the Big 4 audit firms (*BIGN*), change in audit firm (*CHAUD*), busy seasons (*BUSY*), and industry expertise of the auditor (*EXPERT*). All of these control variables are likely to drive the similarity of auditor's reports and KAM disclosure. For instance, *SIZE* is a relevant control as auditors may put additional effort into the creation of auditor's reports for their biggest and thus most prestigious clients, likely resulting in a lower similarity of auditor reporting. On the other hand, auditors are more likely to copy and paste disclosure from the previous year's report to save preparation costs when auditing smaller audit clients.

The only difference in control variables compared to Seebeck and Kaya (2022) is that I do not include *LENGTH* in the basic regressions of H1a and H1b (i.e. examination of year-over-year similarity and within-peer group similarity), because as presented in Appendix C, the denominator used to compute Cosine similarity scores already considers vector lengths and thus normalises the measure to make it invariant to document length (e.g. Hoberg & Phillips, 2010; Loughran & McDonald, 2016).[25] However, when using *BOILERPLATE* as the dependent variable (H1c), I include *LENGTH*, because prior disclosure literature finds a strong positive association between text length and the use of boilerplate language (e.g. Cazier et al., 2021).[26] In line with prior KAM disclosure studies based on UK data, I also do not control for similarity between auditor and management or audit committee disclosure, as conducting large-scale textual analysis of annual reports in the UK is difficult because reports are available in unstructured (i.e. non-machine-readable) formats only (e.g. Burke et al., 2023). All continuous variables are winsorised at the 1% and 99% levels, robust standard errors are clustered by firm, and industry fixed effects are included.[27]

4.3.2 | Capital market reactions (Hypotheses 2 and 3)

While Hypothesis 2 predicts that the similarity of KAM disclosure is not associated with capital market reactions, Hypothesis 3 predicts that the similarity of KAM disclosure might diminish the positive capital market effects from precise KAM disclosure. For the capital market tests, I use the same regression model as used by Seebeck and Kaya (2022) expanded by the test variable *SIMILARITY* (see Equation 1), as well as the interaction term *SPECIFICITY* × *SIMILARITY*, which captures the interplay between similarity and specificity of KAM disclosure.[28] These expansions allow me to examine the direct relationship between similarity and capital market effects as well as potential mitigating effects of disclosure similarity on the positive capital market effects of specificity found in the reference study.

$$\begin{aligned} REACTION_{i,t} = {} & \beta_0 + \beta_1\, SPECIFICITY_{i,t} + \beta_2\, SIMILARITY_{i,t} \\ & + \beta_3\, SPECIFICITY \times SIMILARITY_{i,t} + \beta_4\, REPORT\_DISCL_{i,t} \\ & + \beta_5\, SIZE_{i,t} + \beta_6\, ROA_{i,t} + \beta_7\, MB_{i,t} + \beta_8\, LEV_{i,t} + \beta_9\, CHNI_{i,t} \\ & + \beta_{10}\, STDRET_{i,t} + \beta_{11}\, INCORP_{i,t} + \beta_{12}\, DELAY_{i,t} \\ & + \beta_{13}\, BIGN_{i,t} + FirmFE + TimeFE + \varepsilon_{i,t} \end{aligned} \tag{Eq2}$$

Investors' reactions (*REACTION*) are measured in two different ways: the sum of the 3-day abnormal trading volume around the report filing date (*ABVOL*) and the absolute value of the 3-day abnormal returns around the filing date of the annual report (|*CAR*|). *REPORT_DISCL* is the natural logarithm of the total number of words in the KAM section. T-statistics are calculated using standard errors clustered at the firm level. All continuous variables are winsorised at the 1% and 99% levels to mitigate the effects of outliers.

5 | EMPIRICAL RESULTS

5.1 | Descriptive statistics

Table 2 presents descriptive statistics for the different similarity measures as well as the independent variables used in the main analyses of Hypothesis 1. The mean year-over-year similarity for expanded auditor's reports is 72.1%. The average expanded auditor's report contains 45 boilerplate phrases. Most firms in the sample (98.2%) are audited by the Big 4 audit firms reflecting the high audit market concentration in the UK. Around 16.6% of the financial statements have been audited by industry expert auditors and 9.0% after a change in audit firm. Audit partner changes are much more frequent and affect 24.7% of the firm-year observations, which is in line with mandatory auditor rotation requirements in the UK and the mean tenure of an audit partner of 2.79 years.[29] Finally, I note that the mean length of a KAM section is 1216 words (excluding stop words).

Table 3 shows the tests of overall difference in means of textual similarity measures for the period 1 year before to 6 years after the implementation of ISA 700. The results for the first year post-implementation are in line with findings from Burke et al. (2023) for the US setting, showing that the year-over-year similarity is significantly lower in the first year of post-ISA 700 reporting (52.2%) than in the last year of pre-ISA 700 reporting (84.5%).[30] Thus, auditor reporting in the first year of ISA 700 is less sticky than in the previous years, which is not surprising given that ISA 700 significantly expands the auditor's report. However, it shows that the Cosine similarity measure used in this study appropriately reflects changes in similarity for the given documents.[31] The results further indicate that expanded auditor's reports' stickiness steadily increased over the post-ISA 700 period, from 52.2% in the first year to 77.6% in the sixth year. While there is a statistically significant increase in each of the first 3 years following implementation, it stagnates in the fourth and fifth year, corresponding to the implementation of revised reporting standards ISA 700 and 701 (June 2016). In the sixth year, however, the increase is once more significant and is gradually approaching the pre-ISA 700 level.

Similarly, within-peer group similarity of KAM descriptions and the number of boilerplate phrases used in expanded auditor's reports significantly increased over post-ISA 700 periods. For instance, while

TABLE 2 Descriptive statistics.

Variables	N	Mean	Std. dev.	p25	Median	p75
Dependent variables						
STICKINESS	1045	0.721	0.194	0.602	0.765	0.871
SIM_YEAR	1045	0.500	0.289	0.250	0.500	0.750
SIM_AUDITOR	1045	0.500	0.289	0.250	0.500	0.750
SIM_RISK	1045	0.500	0.289	0.250	0.500	0.750
SIM_PEER	1045	0.500	0.289	0.250	0.500	0.750
BOILERPLATE	1045	45.43	28.30	25	41	59
SIM_ALL	1045	0.500	0.289	0.250	0.500	0.750
Independent variables						
TIMESREP	1045	3.315	1.555	2	3	5
SIZE	1045	14.79	1.470	13.75	14.59	15.63
ROA	1045	7.663	8.321	3.880	6.810	11.18
LEV	1045	3.777	1.855	3.179	3.694	4.273
MB	1045	3.951	5.449	1.653	2.811	4.894
MOWN	1045	0.481	0.500	0	0	1
LOSS	1045	0.107	0.309	0	0	0
CFO	1045	11.02	2.312	10.00	11.02	12.04
USLIST	1045	0.092	0.289	0	0	0
DELAY	1045	60.22	12.74	53	59	67
BIGN	1045	0.982	0.134	1	1	1
CHAUD	1045	0.090	0.286	0	0	0
EXPERT	1045	0.166	0.372	0	0	0
BUSY	1045	0.611	0.488	0	1	1
LENGTH	1045	1216	665	728	996	1479

Notes: This table presents descriptive statistics for the textual similarity measures used as dependent variables and for all independent variables. The descriptive statistics for *LENGTH* are presented as the total number of words for easier interpretation. The regression analyses, however, are based on their natural logarithms. All continuous variables are winsorised at the 1% and 99% levels.

TABLE 3 Univariate time-series analysis.

	Pre-ISA 700 ($n = 180$)		First year ($n = 157$)			Post-ISA 700 Second year ($n = 206$)			Third year ($n = 208$)		
	Mean value [0]	Std. dev.	Mean value [1]	Std. dev.	Mean diff. [1–0]	Mean value [2]	Std. dev.	Mean diff. [2–1]	Mean value [3]	Std. dev.	Mean diff. [3–2]
STICKINESS	0.845	0.208	0.522	0.196	−0.324***	0.716	0.173	0.194***	0.755	0.114	0.040**
SIM_YEAR	- *n/a* -		0.051	0.017	*n/a*	0.052	0.011	0.001*	0.054	0.010	0.002***
SIM_AUDITOR	- *n/a* -		0.055	0.001	*n/a*	0.051	0.002	−0.004**	0.056	0.001	0.005***
SIM_RISK	- *n/a* -		0.060	0.051	*n/a*	0.066	0.018	0.006***	0.067	0.016	0.040
SIM_PEER	- *n/a* -		0.055	0.024	*n/a*	0.058	0.010	0.003**	0.060	0.009	0.002**
BOILERPLATE	- *n/a* -		33.90	23.81	*n/a*	37.74	11.71	3.838**	44.15	14.66	6.410***
			Fourth year ($n = 197$)			Fifth year ($n = 179$)			Sixth year ($n = 98$)		
			Mean value [4]	Std. dev.	Mean diff. [4–3]	Mean value [5]	Std. dev.	Mean diff. [5–4]	Mean value [6]	Std. dev.	Mean diff. [6–5]
STICKINESS			0.784	0.157	0.029	0.769	0.164	−0.016	0.776	0.061	0.007**
SIM_YEAR			0.054	0.010	0.000	0.058	0.013	0.004***	0.063	0.013	0.005***
SIM_AUDITOR			0.058	0.000	0.002**	0.062	0.001	0.004***	0.068	0.001	0.006***
SIM_RISK			0.068	0.014	0.001	0.074	0.019	0.006***	0.081	0.024	0.007***
SIM_PEER			0.061	0.008	0.001*	0.065	0.010	0.004***	0.071	0.013	0.006***
BOILERPLATE			46.57	13.77	2.423*	52.70	18.51	6.128***	65.66	22.93	12.96***

Notes: This table presents the differences in mean values for the textual similarity measures in the year preceding the implementation of ISA 700 and in the first six post-ISA 700 periods. The within-peer group measures (i.e. *SIM_YEAR*, *SIM_AUDITOR* and *SIM_RISK*) are presented as absolute values for informational purposes. The regression analyses as well as the aggregate measure *SIM_PEER*, however, are based on percentile measures. All four measures as well as the number of boilerplate phrases (*BOILERPLATE*) can be computed for KAM disclosure only. Thus, there is no data for these variables in the pre-ISA 700 years. All variables are described in detail in Appendix A.

***Statistical significance at the 1% level.

**Statistical significance at the 5% level.

*Statistical significance at the 10% level.

the mean KAM section included 34 boilerplate four-grams in the first year following the implementation of ISA 700, it continuously increased to 65 in the sixth year, thus having almost doubled. Interestingly, the mean number of KAMs per auditor's report decreases over the same observation period from 3.98 KAMs in year one to about 3.83 KAMs in year six (see Figure 4). This indicates that the average KAM contains significantly more boilerplate phrases in the later post-ISA 700 periods. Together, these findings suggest that expanded auditor's reports converge over time.

5.2 | Correlations

Table 4A presents Pearson correlations among all variables included in the main regression analyses for Hypothesis 1. The results indicate that the number of times reported under ISA 700 (*TIMESREP*) is positively associated with all three measures of textual similarity.[32] The coefficients among the independent variables range between −0.463 (between *LOSS* and *ROA*) and 0.475 (between *LENGTH* and *SIZE*), thus indicating that multicollinearity is no problem. The results of additional variance inflation factor (VIF) collinearity tests conducted for each multivariate regression equation confirm this finding, showing a mean VIF of 2.15 for the main regression analysis of Hypothesis 1 and 2.71 for the main regression analysis of Hypothesis 2 (results are untabulated).

5.3 | Results for Hypothesis 1

5.3.1 | Year-over-year similarity of auditor's reports (H1a)

Table 5A shows results for the regression analyses on auditor's reports' year-over-year similarity (*STICKINESS*) to test H1a. The coefficient for *TIMESREP* is positive and statistically significant for all specifications of *STICKINESS*, suggesting that expanded auditor's reports become stickier over time.[33] This is in line with the existence of strong incentives for auditors to stick to the previous year's auditor's reports and generic disclosure. For instance, minimising changes to the previous year and using boilerplate phrases requires less time and effort when creating the reports. It reduces the time-consuming need for coordination with central reporting departments, which typically exist in bigger audit firms. In addition, recycling KAM disclosure that has already been coordinated with the management in the previous year ensures that

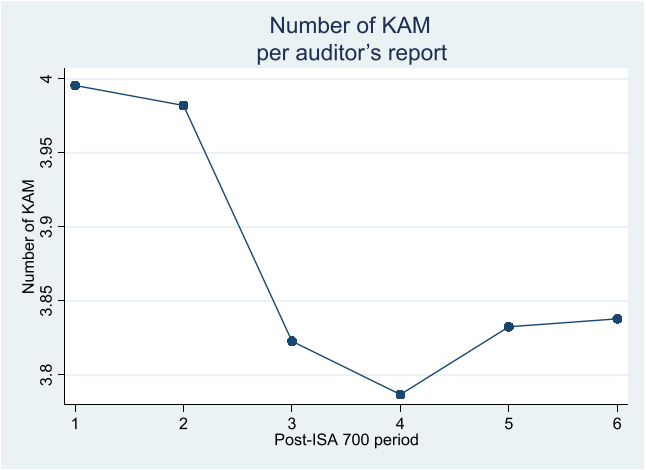

FIGURE 4 Number of KAM per auditors report. Figure 4 shows the mean number of KAMs per expanded auditor's report for post-ISA 700 periods. Figures 2 to 4 cover post-ISA 700 periods only, because KAM disclosure is not available for pre-ISA 700 periods. In 2016, revised auditor reporting standards ISA 700 and 701 (June 2016) were implemented, resulting in stagnating Cosine similarity scores in post-ISA 700 periods four and five. [Colour figure can be viewed at wileyonlinelibrary.com]

proprietary information remains protected, as requested by the management (Bernard et al., 2018).[34] Overall, my results suggest that the stickiness of expanded auditor's reports increases over time and thus do not support H1a, which was proposed in the null form.

I further note that many of the firm and auditor controls are significant, suggesting that the regression model is well-specified and the empirical design is powerful.[35] For instance, the occurrences of a loss (*LOSS*), a change in audit firm (*CHAUD*) and audit work conducted during busy seasons (*BUSY*) are negatively associated with year-over-year textual similarity across all model variations. The observation that auditor reporting is less sticky during the busy season might be counterintuitive, as one might expect higher workloads to lead to poorer reporting quality because of time constraints. However, during the busy season, auditors generally complete multiple engagements in parallel. In addition, they often work closely in teams and share knowledge and experiences with one another across engagement teams. When auditors share knowledge and experiences, they might be influenced by insights that their peers have discovered. If these insights are deemed valuable and relevant, multiple auditors might integrate them into their KAM reporting. As a result, their auditor's reports become more similar to each other and less sticky. This reasoning is in line with the higher within-peer group similarity across audit clients during busy season presented in Table 5B.

5.3.2 | Within-peer group similarity of KAM descriptions (H1b)

Table 5B presents coefficients and *T*-statistics for the regression analyses on within-peer group similarity to test H1b. The dependent variable is the textual similarity of a given KAM description to all KAM descriptions in the same peer group, as defined in Section 4.2.2. (i.e. *SIM_YEAR*, *SIM_AUDITOR* and *SIM_RISK*), as well as an aggregate measure of the three variables (*SIM_PEER*). Similar to the findings for year-over-year similarity, I find that the coefficient of *TIMESREP* is significantly positive for all specifications, suggesting that auditors increasingly coordinate the disclosure in KAM descriptions with peer reports and thus use more similar language over time. The positive relations between *BIGN* and most of the similarity measures analysed in Tables 5B and 5C indicate that the Big 4 auditors tend to use more similar and boilerplate language. This is potentially because of the fact that the Big 4 auditors have more standardised reporting processes and were notably very fast in creating large databases of mostly industry and risk-type-specific standard wording templates for KAM descriptions.[36]

5.3.3 | Boilerplate language in KAM sections (H1c)

Table 5C presents results for the regression on the number of boilerplate phrases in KAM sections (H1c). The results show a positive and statistically significant coefficient for *TIMESREP*, indicating that, on average, there is an annual increase of six additional boilerplate phrases in the KAM sections. This finding is robust to adjustments in the frequency limits which determine when an n-gram becomes a boilerplate phrase and also remain stable when amending the number of words which build an n-gram (i.e. four-gram, five-gram, and six-gram).

Interestingly, the results indicate that *LOSS* has a positive impact on the use of boilerplate language as well as on within-peer group similarity but a negative impact on stickiness. There are several reasons why auditor reporting may appear less sticky when an audit client reports a loss: First, auditors may exercise greater caution and diligence when examining financial statements of a client with losses, as the likelihood of misstatements or fraud can be higher in such situations. This increased scrutiny could result in more detailed and less sticky KAM reporting. Second, firms reporting losses are likely to face more significant pressure from investors, creditors, and regulators. Consequently, audit committees may be even more inclined than in the previous years to ensure that auditor reporting is accurate and transparent to maintain credibility, which can lead to less stickiness in reporting. Third, firms with losses are more likely to attract regulatory attention, as financial difficulties may indicate underlying issues or potential misstatements. This increased regulatory scrutiny could prompt both audit committees and auditors to be more careful in their reporting and auditing, resulting in less stickiness. However, audit clients reporting a loss can also result in more similar and boilerplate KAM disclosure. For instance, by using generic and boilerplate language, auditors might attempt to limit their exposure to legal liabilities or regulatory actions. In addition, auditors may put stronger emphasis on compliance with the standards, as with increased scrutiny on clients reporting losses, auditors may focus on ensuring that the auditor's report complies with all relevant standards and requirements. This might lead them to rely more on standardised boilerplate

TABLE 4 Pearson correlations between dependent variables and control variables.

	Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	STICKINESS	1.000																	
2	SIM_PEER	0.068	1.000																
		0.029																	
3	BOILERPLATE	***0.098***	***0.349***	1.000															
		0.002	0.000																
4	TIMESREP	***0.473***	***0.159***	***0.302***	1.000														
		0.000	0.000	0.000															
5	SIZE	0.010	−***0.167***	***0.103***	***0.107***	1.000													
		0.740	0.000	0.001	0.001														
6	ROA	0.023	−0.059	−***0.140***	−0.045	−***0.266***	1.000												
		0.463	0.057	0.000	0.148	0.000													
7	LEV	0.015	−***0.121***	***0.113***	0.041	***0.283***	−***0.139***	1.000											
		0.619	0.000	0.000	0.185	0.000	0.000												
8	MB	−***0.031***	−0.035	−0.048	−0.039	−***0.179***	***0.364***	−0.022	1.000										
		0.002	0.259	0.119	0.212	0.000	0.000	0.486											
9	MOWN	−0.004	0.012	−0.008	−0.014	−***0.389***	−***0.088***	−0.034	0.028	1.000									
		0.889	0.709	0.806	0.651	0.000	0.004	0.267	0.368										
10	LOSS	0.077	0.025	***0.087***	−0.008	0.078	−***0.463***	***0.128***	−***0.082***	***0.087***	1.000								
		0.012	0.420	0.005	0.784	0.012	0.000	0.000	0.008	0.005									
11	CFO	−0.055	−***0.135***	−***0.082***	−0.068	−***0.164***	***0.161***	−0.042	***0.123***	0.068	0.032	1.000							
		0.076	0.000	0.008	0.029	0.000	0.000	0.176	0.000	0.029	0.302								
12	USLIST	−0.005	−***0.195***	0.009	0.025	***0.537***	0.026	***0.148***	−0.012	−***0.267***	−0.025	−0.012	1.000						
		0.864	0.000	0.763	0.418	0.000	0.407	0.000	0.700	0.000	0.429	0.699							
13	BIGN	0.030	***0.105***	***0.096***	0.005	0.038	−0.039	0.056	0.051	−0.070	0.024	−0.052	0.043	1.000					
		0.338	0.001	0.002	0.884	0.217	0.210	0.071	0.101	0.025	0.438	0.094	0.162						
14	DELAY	−***0.031***	−0.058	−0.052	−0.003	−***0.121***	−***0.086***	−***0.086***	−0.053	0.005	***0.081***	***0.101***	0.020	−0.058	1.000				
		0.000	0.059	0.093	0.927	0.000	0.005	0.005	0.087	0.867	0.009	0.001	0.520	0.062					
15	CHAUD	−***0.315***	0.045	0.053	***0.108***	0.020	−0.007	−0.001	−0.042	−0.015	0.010	0.002	0.016	−0.007	0.058	1.000			
		0.000	0.149	0.085	0.000	0.521	0.809	0.984	0.175	0.627	0.747	0.953	0.610	0.814	0.060				
16	EXPERT	***0.026***	0.057	−0.006	−0.039	0.000	−0.019	−***0.089***	−0.025	−0.012	−0.046	−0.028	***0.108***	0.061	0.006	0.040	1.000		
		0.004	0.067	0.857	0.209	0.994	0.540	0.004	0.418	0.705	0.136	0.371	0.001	0.050	0.843	0.197			
17	BUSY	−0.025	***0.075***	0.070	0.063	0.053	−0.046	−0.034	−0.033	−0.038	***0.092***	−0.016	***0.090***	0.009	***0.167***	−0.031	−0.015	1.000	
		0.416	0.006	0.024	0.042	0.085	0.141	0.268	0.293	0.224	0.003	0.597	0.004	0.769	0.000	0.321	0.635		
18	LENGTH	0.113	0.192	***0.371***	***0.155***	***0.475***	−0.009	0.273	−0.090	−0.146	***0.153***	−0.076	0.282	***0.064***	0.000	0.040	0.015	−0.037	1.000
		0.237	0.625	0.000	0.002	0.000	0.621	0.446	0.134	0.240	0.000	0.114	0.300	0.001	0.242	0.192	0.637	0.230	

Notes: This table presents the Pearson's correlations results for *STICKINESS*, *SIM_PEER* and *BOILERPLATE*, as well as the control variables used in the main analyses. Correlations significant at the 1% level are bolded and italicised. All variables are described in detail in Appendix A.

TABLE 5 Results for Hypothesis 1.

Panel A: year-over-year similarity of expanded auditor's reports.					
Variables	**Exp. sign**	**(1) STICKINESS**	**(2) $STICKINESS_{TF}$**	**(3) $STICKINESS_{SW}$**	**(4) $STICKINESS_{TF\_SW}$**
TIMESREP	+	**0.073*** **	**0.096*** **	**0.074*** **	**0.100*** **
		(14.537)	**(21.503)**	**(14.855)**	**(22.219)**
SIZE	−	−0.021**	−0.011	−0.018**	−0.007
		(−2.462)	(−1.418)	(−2.166)	(−0.880)
ROA	−	0.001	−0.001	0.001	−0.001
		(0.595)	(−0.430)	(0.537)	(−1.094)
LEV	+	0.000	0.000	−0.000	0.000
		(0.236)	(1.068)	(−0.040)	(1.414)
MB	+	0.000	0.000	0.000	0.000
		(0.807)	(1.524)	(0.936)	(1.579)
MOWN	−	−0.032*	−0.004	−0.030	−0.009
		(−1.737)	(−0.206)	(−1.609)	(−0.525)
LOSS	−	−0.068**	−0.068**	−0.080**	−0.059**
		(−2.137)	(−2.406)	(−2.519)	(−2.182)
CFO	−	−0.454*	−0.311	−0.431	−0.241
		(−1.702)	(−1.332)	(−1.611)	(−1.016)
USLIST	?	0.101***	0.020	0.113***	0.012
		(2.594)	(0.631)	(2.933)	(0.374)
DELAY	−	−0.000	−0.000	−0.000	−0.000
		(−0.628)	(−0.570)	(−0.473)	(−0.574)
BIGN	+	−0.005	0.046	−0.004	0.019
		(−0.088)	(0.880)	(−0.069)	(0.348)
CHAUD	−	−0.284***	−0.371***	−0.284***	−0.345***
		(−13.925)	(−22.914)	(−13.752)	(−21.277)
EXPERT	+	0.021*	0.002	0.015	0.001
		(1.733)	(0.094)	(0.548)	(0.054)
BUSY	?	−0.057***	−0.051***	−0.058***	−0.044**
		(−3.079)	(−2.952)	(−3.165)	(−2.529)
Constant		0.658***	0.387***	0.606***	0.356**
		(4.011)	(2.592)	(3.706)	(2.409)
Observations		1045	1045	1045	1045
Industry FE		Yes	Yes	Yes	Yes
Adj. *R*-squared		0.268	0.386	0.278	0.388

Panel B: within-peer group similarity of KAM descriptions.				
Variables	**(5) SIM_YEAR**	**(6) SIM_AUDITOR**	**(7) SIM_RISK**	**(8) SIM_PEER**
TIMESREP	**0.030*** **	**0.040*** **	**0.011***	**0.028** **
	(5.376)	**(6.827)**	**(1.957)**	**(2.400)**
SIZE	−0.019**	−0.014	0.001	−0.012*
	(−2.031)	(−1.495)	(0.077)	(−1.847)
ROA	−0.005***	−0.003**	−0.002*	−0.003**
	(−3.373)	(−2.186)	(−1.699)	(−2.374)
LEV	−0.000	−0.000	−0.000	−0.000
	(−1.168)	(−0.838)	(−1.182)	(−1.024)

(Continues)

TABLE 5 (Continued)

Panel B: within-peer group similarity of KAM descriptions.				
Variables	(5) SIM_YEAR	(6) SIM_AUDITOR	(7) SIM_RISK	(8) SIM_PEER
MB	0.000	0.000	0.000	0.000
	(0.316)	(0.619)	(1.622)	(1.138)
MOWN	0.001	−0.001	0.009	0.004
	(0.034)	(−0.048)	(0.488)	(0.251)
LOSS	0.038*	0.036*	0.011	0.028
	(1.597)	(1.739)	(0.363)	(1.177)
CFO	−0.800***	−0.896***	−0.900***	−0.898***
	(−3.212)	(−3.346)	(−3.523)	(−3.316)
USLIST	−0.068*	−0.070*	−0.014	−0.063*
	(−1.892)	(−1.894)	(−0.353)	(−1.900)
DELAY	0.001	0.002**	−0.000	0.001*
	(1.319)	(2.441)	(−0.300)	(1.959)
BIGN	0.169***	0.046	0.150***	0.110***
	(3.271)	(0.719)	(3.269)	(3.274)
CHAUD	0.014	0.014	0.031	0.028
	(0.486)	(0.488)	(1.107)	(0.762)
EXPERT	−0.022	−0.022	−0.020	−0.022
	(−0.885)	(−0.872)	(−0.795)	(−0.821)
BUSY	0.026	0.010	0.056***	0.040**
	(1.401)	(0.527)	(3.079)	(2.439)
Constant	0.621***	0.582***	0.402**	0.539***
	(3.854)	(3.527)	(2.510)	(3.322)
Observations	1045	1045	1045	1045
Industry FE	Yes	Yes	Yes	Yes
Adj. *R*-squared	0.269	0.225	0.268	0.255
Panel C: number of boilerplate phrases per KAM section.				
Variables	(9) BOILERPLATE 4-grams	(10) BOILERPLATE 5-grams	(11) BOILERPLATE 6-grams	(12) SIM_ALL
TIMESREP	**6.047*** **	**6.190*** **	**5.832*** **	**0.033*** **
	(11.260)	**(7.439)**	**(5.437)**	**(4.170)**
SIZE	−1.364*	−0.491	0.021	−0.226
	(−1.697)	(−0.892)	(0.329)	(−1.040)
ROA	−0.194	−0.091*	0.123	−0.002*
	(−1.601)	(−1.782)	(0.081)	(−1.689)
LEV	0.023	0.026	0.029	−0.000
	(0.871)	(0.865)	(0.829)	(−1.002)
MB	0.000	0.000	0.001	0.000
	(0.181)	(0.148)	(0.184)	(0.989)
MOWN	0.733	0.861	0.565	0.052
	(0.425)	(0.798)	(0.358)	(0.341)
LOSS	3.415*	3.576*	2.924*	0.182*
	(1.837)	(1.699)	(1.739)	(1.727)
CFO	1.671	1.861	1.148	−0.487**
	(0.074)	(0.036)	(0.153)	(−2.306)

TABLE 5 (Continued)

Panel C: number of boilerplate phrases per KAM section.				
Variables	(9) BOILERPLATE 4-grams	(10) BOILERPLATE 5-grams	(11) BOILERPLATE 6-grams	(12) SIM_ALL
USLIST	−1.804	−1.048	−1.057	−0.060*
	(−0.497)	(−0.552)	(−0.581)	(−1.825)
DELAY	−0.035	−0.071	−0.066	0.001*
	(−0.559)	(−0.581)	(−0.560)	(1.959)
BIGN	24.665***	18.010***	16.932***	0.293***
	(8.880)	(6.923)	(6.276)	(3.753)
CHAUD	−0.537	0.290*	0.111*	−0.038*
	(−0.195)	(1.903)	(1.820)	(−1.751)
EXPERT	−0.214	−0.256	0.184	0.114
	(−0.089)	(−0.075)	(0.023)	(0.365)
BUSY	3.163*	3.301**	2.880**	0.034**
	(1.926)	(2.229)	(2.035)	(2.559)
LENGTH	8.579***	8.888***	8.097***	0.165
	(13.771)	(10.912)	(8.192)	(1.087)
Constant	−15.064	−13.928	−10.347*	0.709***
	(−1.061)	(−1.327)	(−1.724)	(3.800)
Observations	1045	1045	1045	1045
Industry FE	Yes	Yes	Yes	Yes
Adj. *R*-squared	0.368	0.381	0.383	0.304
Panel D: categorical time effects model.				
Variables	(13) STICKINESS	(14) SIM_PEER	(15) BOILERPLATE	(16) SIM_ALL
Year 2	0.065***	0.009	2.154***	0.023**
	(4.872)	(0.497)	(5.695)	(2.610)
Year 3	0.043**	0.009*	3.925***	0.020**
	(2.226)	(1.748)	(5.163)	(2.796)
Year 4	0.030	0.011*	2.001**	0.017*
	(1.011)	(1.866)	(2.136)	(1.704)
Year 5	0.024	0.025***	4.971***	0.013**
	(1.084)	(3.795)	(4.345)	(2.328)
Year 6	0.026*	0.029***	3.892***	0.014**
	(1.897)	(3.469)	(7.494)	(2.431)
Constant	0.600***	0.497***	−9.101	0.620***
	(3.916)	(3.683)	(−0.784)	(3.994)
Observations	1045	1045	1045	1045
Industry FE	Yes	Yes	Yes	Yes
Adj. *R*-squared	0.229	0.231	0.389	0.286

Notes: This table reports the coefficients for the multivariate regression on different similarity variables. The dependent variables are various specifications of year-over-year similarity for the full auditor's report in Panel A, of within-peer group similarity of KAM descriptions in Panel B, and of boilerplate language used in KAM sections in Panel C. Panel D shows the results of a Categorical Time Effects Model in which TIMESREP is replaced with year dummies (e.g. year 2 for the second post-ISA 700 period). The last column in Panel C reports the coefficients for an aggregate similarity measure. In the column headings of Panel A, 'TF' indicates that the TF weighting scheme was used, and 'SW' indicates that stop words are included. In the column headings of Panel C, 'n-grams' specifies the number of words that form an n-gram. *T*-statistics are shown in parentheses. All variables are described in detail in Appendix A.

***Statistical significance at the 1% level.

**Statistical significance at the 5% level.

*Statistical significance at the 10% level.

TABLE 6 Results for Hypothesis 2.

Panel A: capital market consequences of similarity and abnormal similarity in auditor's reports.						
Variables	**Exp. sign**	**(17) ABVOL**	**(18) ABVOL**	**Exp. sign**	**(19) |CAR|**	**(20) |CAR|**
SIM_ALL	**?**	**−0.003**		**?**	**−0.012**	
		(−1.029)			**(−1.111)**	
ABN_SIM_ALL	**?**		**−0.058***	**?**		**−0.060**
			(−1.770)			**(−0.618)**
REPORT_DISCL	−	−0.042	−0.029	−	−0.024	−0.037
		(−0.120)	(−0.231)		(−0.523)	(−0.359)
SIZE	−	−0.488**	−0.482**	−	−0.454	−0.500
		(−2.773)	(−2.611)		(−0.350)	(−0.976)
ROA	?	−0.321	−0.249	+	0.405**	0.510**
		(−0.323)	(−0.332)		(2.088)	(2.798)
MB	+	0.015	0.105	?	0.001	0.001
		(0.485)	(0.548)		(1.133)	(1.000)
LEV	?	0.014**	0.102**	−	−0.007*	−0.005*
		(2.432)	(2.524)		(−1.861)	(−1.728)
CHNI	+	0.066	0.269	+	0.014*	0.012*
		(0.082)	(0.216)		(1.744)	(1.735)
STDRET	+	−2.085*	−2.075	+	1.465***	1.439***
		(−1.676)	(−0.276)		(3.596)	(3.740)
INCORP	+	0.282***	0.274***	+	0.009*	0.010*
		(2.631)	(2.927)		(1.707)	(1.731)
DELAY	?	−0.022*	−0.402	?	0.015	0.009*
		(−1.906)	(−0.656)		(0.400)	(1.859)
BIGN	+	0.692**	0.693**	?	−0.041*	−0.037*
		(2.329)	(2.406)		(−1.703)	(−1.711)
Constant		−6.491	−6.534		−0.100	−0.211
		(−1.299)	(−1.208)		(−0.993)	(−0.376)
Observations		1045	1045		1045	1045
Firm FE		Yes	Yes		Yes	Yes
Time FE		Yes	Yes		Yes	Yes
Adj. *R*-squared		0.062	0.064		0.099	0.093

Panel B: moderating effect of similarity on capital market effects from KAM specificity.				
Variables	**(21) ABVOL**	**(22) ABVOL**	**(23) |CAR|**	**(24) |CAR|**
SPECIFICITY	0.830***	0.805***	0.026***	0.031***
	(4.701)	(4.555)	(4.878)	(5.200)
SIM_ALL	−0.005*	−0.003	−0.009	−0.018*
	(−1.771)	(−0.928)	(−1.067)	(−1.823)
SPECIFICITY × SIM_ALL		**−0.201*** **		**−0.022****
		(−3.630)		**(−1.790)**
Observations	1045	1045	1045	1045
Controls	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Adj. *R*-squared	0.172	0.183	0.129	0.110

TABLE 6 (Continued)

Panel C: moderating effect of abnormal similarity on capital market effects from KAM specificity.				
Variables	**(25) ABVOL**	**(26) ABVOL**	**(27) |CAR|**	**(28) |CAR|**
SPECIFICITY	0.873***	0.916***	0.034***	0.040***
	(4.537)	(4.503)	(3.855)	(4.219)
ABN_SIM_ALL	−0.033	−0.050*	−0.048*	−0.055
	(−0.710)	(−1.609)	(−1.713)	(−0.404)
SPECIFICITY × ABN_SIM_ALL		**−0.309*** **		**−0.026*** **
		(−4.530)		**(−3.974)**
Observations	1045	1045	1045	1045
Controls	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Adj. *R*-squared	0.110	0.188	0.192	0.174

Notes: This table reports the coefficients for the multivariate regression analysis of capital market reactions including robust *T*-statistics in parentheses. The dependent variables are abnormal trading volume (*ABVOL*) and cumulative abnormal returns (|*CAR*|), as indicated in the column headings. Panel A shows the results of estimating equity market effects including (abnormal) similarity of the auditor's report. Panels B and C add in *SPECIFICTY* of the KAM descriptions, as well as interaction variables for *SPECIFICITY* x (abnormal) *SIMILARITY*. All variables are described in detail in Appendix A.
***Statistical significance at the 1% level.
**Statistical significance at the 5% level.
*Statistical significance at the 10% level.

language. Finally, by using generic and boilerplate language, auditors might aim to avoid drawing attention to the client's financial losses or provoking potential disputes with the client. In line with this reasoning, ISA 701.14 states that the description of KAMs should not reveal information that is confidential or proprietary. Auditors can strategically avoid such conflicts by aligning the wording with the standard.

The last column in Table 5C shows the results for the regression on an aggregate similarity measure (*SIM_ALL*), which is used in subsequent capital market tests. It includes *LENGTH* as a control. The insignificant coefficient for *LENGTH* is representative for all regression analyses to test for H1a and H1b. However, coefficients for text length are consistently positive for all regressions on *BOILERPLATE* and therefore included when testing for H1c.

5.3.4 | Further discussion of the results for Hypothesis 1

The positive association between *TIMESREP* and *SIMILARITY* holds for all specifications of similarity and is stable even against the background of increasingly longer texts. Together, the findings indicate that expanded auditor reporting converges over time. This is consistent with the existence of strong institutional and environmental incentives for auditors to stick to the previous year's reporting, coordinate KAM disclosures with peer reports, or to use boilerplate language.

However, one must exercise caution when drawing conclusions about potential consequences and reasons. On the one hand, sticky, generic and boilerplate disclosure bears the risk of not providing any useful information for financial statement users. On the other hand, high KAM disclosure similarity is not necessarily negative, because some potentially relevant information on persisting matters must sensibly be carried forward from year to year. Removing this information from auditor's reports would result in a loss of reporting quality. Moreover, continuity in reporting can increase comparability and thus readability (Loughran & McDonald, 2014). Similarly, notwithstanding the fact that the use of similar language in peers' KAM descriptions generally indicates rather generic reporting across audit clients, it can be driven by similar circumstances and risk environments faced by the auditors during the audit and thus, be reasonable.

Finally, the increasing use of boilerplate phrases over time, which clearly indicates a shift towards more generic reporting, can also be driven by relevant formulations that have proven particularly useful over time. These are consequently used more frequently in later years as the language of expanded auditor's reports likely becomes better rehearsed over time.

In order to better understand the potential consequences of increasing KAM disclosure similarity, I test for potential equity market consequences in the examinations of Hypotheses 2 and 3.

5.4 | Results for Hypotheses 2 and 3

I run a series of regression analyses to examine potential capital market consequences of similarity in expanded auditor's reports. However, as presented in Table 6A, I do not find consistent evidence of a direct association between any of the similarity measures described in Section 4.2. and cumulative abnormal returns (|*CAR*|) or abnormal

trading volume (*ABVOL*). Furthermore, I test for an association between abnormal similarity and capital market effects. To this end, I decompose the different similarity measures into normal and abnormal similarity components. The expected similarity for a given auditor's report is determined based on Equation (1). The abnormal similarity is the difference between the observed similarity values and the estimated values. Thus, it surrogates for auditors' intended or unintended extraordinary choice of sticky or generic reporting. I do not find consistent results either. I continue to find inconsistent results when amending the technical specifications of the similarity measures, as well as estimation periods, event windows and weighting of returns for capital market variables (results are untabulated). The adjusted *R*-squared is <0.1 for all model specifications presented in Table 6A, indicating that they have low explanatory power for capital market reactions.

Next, I examine the interplay between similarity of KAM disclosure and specificity. While specificity is a direct measure of information content which goes beyond pure writing style, similarity is a powerful measure to compare different texts. Given that Seebeck and Kaya (2022) find an association between more precise descriptions of KAMs and equity market reactions and Deneuve et al. (2023) find that risk descriptions of KAMs are informative if they provide new information compared to industry peers and the previous year simultaneously, I examine if similarity moderates this effect. The positive coefficients for *SPECIFICITY* in Tables 6B and 6C show that the positive association between specificity and equity market effects also holds for this study. However, the negative coefficient for the interaction term of *SPECIFICITY* and (abnormal) similarity indicates that (abnormal) similarity diminishes the positive effect of specificity on capital market effects. These findings suggest that the positive capital market effects of precise information in KAMs are driven by client- and engagement-specific risk disclosures. The adjusted *R*-squares are in line with prior studies examining the effect of specific disclosure on capital markets (e.g. Hope et al., 2016; Seebeck & Kaya, 2022).

6 | CONCLUSION

This study is among the first to examine the development of expanded auditor reporting over time. I find consistent and robust evidence for growing year-over-year similarity (stickiness), within-peer group similarity (generality) and boilerplate language. Moreover, I provide initial evidence of a moderating effect that textual similarity of KAM disclosure has on capital market effects. Together, these findings suggest that auditors have successively developed standard text modules for KAM disclosures that undermine the informative value of expanded auditor's reports and add to the auditing literature on the fundamental question of how KAMs are worded.

My findings can inform regulators about some potential unintended consequences of expanding the auditor's report and of mandating principle-based standards. By employing textual analyses on expanded auditor's reports, regulators can identify instances of sticky, generic and boilerplate disclosures. This, in turn, equips them to take several specific actions to counteract these issues such as conducting regular training sessions and workshops for auditors to reinforce the importance of entity- and reporting period-specific reporting in KAMs, implementing review and feedback mechanisms, and devising reporting standards and guidelines that foster useful KAM disclosure and counteract disclosure overload through the limitation of non-client-specific legalistic and technical language. However, my study only provides initial evidence on the long-term development of expanded auditor reporting in the UK. One must exercise caution when transferring these results to other countries as there may be technical differences between the standards (Minutti-Meza, 2021) as well as differences in how auditors embrace the new standards (Francis, 2019). For instance, initial evidence from the United States shows that on average auditors disclose only 1–2 CAMs per report, compared to almost 4 KAMs in the UK. In addition, an important difference between the UK and the United States is that the British auditors reveal the name of the engagement partner in their reports. Consequently, UK audit partners may be even more concerned about language used in their reports.

Overall, my study provides initial evidence on the long-term development of expanded auditor reporting and its consequences for financial statement users. Importantly, it only investigates short-term capital market reactions. Further evidence in this largely unexplored field is therefore needed. In particular, I call for more research on the association between the development of textual characteristics of expanded auditor's reports and audit quality and delay.

ACKNOWLEDGEMENTS

I thank the participants at the ACA Research Symposium St. Gallen 2017, EARNet Symposium 2017, EAA Annual Congress 2018 and 2022, Constructor University, University of Erlangen-Nuremberg, as well as Klaus Henselmann, Devrimi Kaya, Nicolas Schwarz, Tami Dinh, Ilias Basioudis, Andrew Ferguson, Lasse Niemi, Tom Rose, Jochen Bigus, Jochen Hundsdoerfer, Andreas Löffler Tobias Svanström (discussant), and David Hay (editor) for their helpful comments. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

The author reports there are no competing interests to declare.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ETHICS STATEMENT

Not applicable because the paper uses archival data from a public domain and does not state or imply any personal information.

ORCID

Andreas Seebeck https://orcid.org/0000-0003-3903-1157

ENDNOTES

[1] ISA 700 from June 2013 initially used the terminology 'risks of material misstatements' (ISA 700.19A), but the revised standards ISA 700 and ISA 701 from June 2016 adopted the term 'key audit matters' from the IAASB (FRC, 2016a, 2016b). The FRC's RMM and KAM reporting requirements are broadly equivalent to the KAM requirements of ISA 701 (IAASB, 2015) and the 'critical audit matters' (CAM) requirements of AS 3101 (PCAOB 2017).

[2] While the IAASB expects that the language used by auditors in the description of a key audit matter '*relates the matter directly to the specific circumstances of the entity, while avoiding generic or standardized language*' (IAASB, ISA 701.A47), the PCAOB requires that '*communication [is] tailored to the audit to avoid standardized language and to reflect the specific circumstances of the matter*' (PCAOB, AS 3101, p. 30).

[3] Possibly, two KAM descriptions can be totally different while having the same readability, tone, specificity, and number of visual aids.

[4] All similarity measures are described in detail in Section 4.2 and in the Appendix.

[5] Cosine similarity measures text similarity by transforming each text into a high-dimensional vector, where each dimension represents a unique word from the combined texts' vocabulary. It then computes the cosine of the angle between these vectors, which essentially reflects the overlap in their content. The closer the cosine value is to 1, the more similar the two texts are in terms of their lexical content, regardless of their word order or length. Appendix C describes the Cosine similarity approach and technical specifications in more detail.

[6] Section 4.2 and Appendix B describe the n-gram approach used to measure boilerplate language in more detail.

[7] Expanded auditor's reports in France including justification of assessments (JOAs) exist since 2003. However, JOAs are different from KAMs because their identification requirements are arguably very general when compared to the FRC's, IAASB's, and PCAOB's standards for KAM (CAM) reporting (Bédard et al., 2014).

[8] ISA 701.A34 requires the auditor to provide '*useful information in a concise and understandable form, while not inappropriately being the provider of original information about the entity*.'

[9] For instance, ISA (UK & Ireland) 701 provides specific guidance on how to judge whether fraud risks are KAMs. Specifically, ISA 701.A21 states that '*depending on their nature, these [fraud] risks may not require significant auditor attention, and therefore would not be considered in the auditor's determination of key audit matters*'. As a result, the number of KAMs which address fraud in revenue recognition dropped significantly after the implementation of ISA 701.

[10] Outside of the UK, many regulators and standard setters in established and emerging economies have recently implemented similar expanded auditor reporting standards (e.g. Auditing and Assurance Standards Board [AASB], 2017; Chinese Institute of Certified Public Accountants [CICPA], 2016; IAASB, 2015; PCAOB, 2017). Although different terminologies and definitions are used, there is significant overlap among the expanded reporting requirements of material audit matters.

[11] Minutti-Mezza (2021), who examines expanded auditor's reports in various jurisdictions, provides a detailed overview.

[12] This observation period rules out confounding factors from the COVID-19 pandemic. Many regulators including the FRC published guidelines for COVID-19 related issues, including KAM reporting (e.g. FRC, 2020; IAASB 2020), that resulted in a high number of similar COVID-19-related KAMs in more recent reports.

[13] Documents are cut straight after the date of the opinion to separate the auditor's reports from the annual reports. This procedure excludes any footnotes to the auditor's report.

[14] The scope of ISA 701 (June 2016) is broader than the scope of ISA 700 (June 2013). However, to use the widest possible observation period while keeping fluctuation of sample firms low and allowing for high data availability of control variables, I restrict the sample to FTSE 350 indexed firms which account for approximately 95% of all listed companies in the UK.

[15] I acknowledge that the primary focus of prior literature on expanded auditor's reports has predominantly centred on KAM-sections rather than the whole auditor's reports. Consequently, my findings on the stickiness of reporting and its implications should be interpreted with caution and may not directly align with the more targeted insights specific to KAM disclosures.

[16] In line with the ICAEW (2017), I identified 20 risk types reported in KAMs, indicating the diverse content of the KAMs. Among the most reported risk types are 'revenue recognition', 'goodwill' and 'intangibles', along with 'tax'.

[17] For example, the mean similarity is 0.5 for a KAM section that contains the two KAMs A and B, where A has a similarity of 0.4 and B has a similarity of 0.6. The similarity of each KAM is derived by comparing the KAM description with all other KAM descriptions that have been reported in the same period (SIM_YEAR), been reported by the same audit firm in the same reporting period (SIM_AUDITOR), or describe the same type of risk for the same reporting period (SIM_RISK).

[18] Because of KAM disclosures being significantly shorter than 10-K disclosures examined by Lang and Stice-Lawrence (2015), I adjusted the frequency limits that determine when a four-gram becomes a boilerplate text. Specifically, only n-grams that occur in at least 50% of the sample firms' KAM sections are considered boilerplate phrases. The results remain robust when using a threshold of 60% (Lang & Stice-Lawrence, 2015) and 75% (Dyer et al., 2023) and dropping common innocuous phrases (e.g. 'as a result of').

[19] The use of four-grams strives a good balance between high specificity (i.e., reduced noise) and context preservation on the one hand, and a feasible requirement of computational resources on the other hand. In additional analyses, I use n-grams of five and six words. Furthermore, I scale the number of boilerplates by document length. The results remain qualitatively the same.

[20] This is an amended regression model of Gutierrez et al. (2018), which is tailored to textual characteristics of expanded auditor's reports. All results remain qualitatively the same when using Tobit regression for *STICKINESS*, *SIM_PEER*, and *SIM_RISK* and Poisson regression for *BOILERPLATE*.

[21] SIM_ALL offers a comprehensive perspective on the similarity of KAMs by capturing insights from all three dimensions.

[22] My results remain qualitatively the same when using principal component analysis for aggregation.

[23] In additional analyses, I use interval measures for the different similarity variables and continue to find similar results.

[24] I acknowledge that there may be potentially relevant variables that have not been included in the model because of data limitations, multicollinearity issues, and the fact that not all potentially relevant factors can be quantified. For instance, because of multicollinearity issues, I could not include both audit firm tenure and audit firm change. As I predict and find a strong association between auditor changes and the stickiness of auditor reporting, I decided to include audit firm change in the regression model at the expense of auditor tenure. Moreover, third-party data availability on audit committee characteristics for the sample firms is poor, and conducting large-scale textual analysis of annual reports in the UK is difficult because reports are only available in unstructured (i.e., non-machine-readable) formats. Thus, I did not include them as controls.

[25] In untabulated robustness checks, I find that the results remain qualitatively the same when including *LENGTH* as a control variable. In additional robustness checks, I follow Brown and Tucker (2011), who use the Taylor expansion to control for a potential mechanical relationship between similarity and document length (i.e., the longer a pair of documents, the more likely a word is included in both documents). Specifically, I empirically approximate the relationship between similarity and auditor's report length by a Taylor expansion at 0 and calculate the expected difference score given the auditor's report length. Next, I compute the adjusted similarity score as the similarity score minus this expected score. My results remain qualitatively unchanged.

[26] In additional analyses, I scale *BOILERPLATE* by document length and drop *LENGTH* as a control. My results remain qualitatively unchanged.

[27] I ran all the main tests using an alternative auditor fixed effects model (excluding *BIGN*). The results and inferences are consistent with the main results.

[28] All measures are described in Appendix A.

[29] The maximum tenure of an audit partner is 5 years, as regulated by the Auditing Practices Board's Ethical Standard 3 (FRC, 2009).

[30] Using the TF weighting scheme, the mean year-over-year similarity of auditor's reports decreases from 98.1% to 80.4% because of the implementation of ISA 700.

[31] As there have not been significant changes in auditor reporting requirements in the pre-ISA 700 period, the high textual similarity of the auditor's reports issued in the year immediately preceding the implementation of ISA 700 can be regarded as representative for all pre-ISA 700 years.

[32] There is a significant positive relation between all three within-peer group similarity variables and the number of times reported.

[33] In robustness checks, I used a Categorical Time Effects Model including year dummies for post-ISA 700 periods 2–6. To avoid multicollinearity, the dummy variable for the omitted first year serves as the reference category. The results presented in Table 5D confirm the findings of the Linear Trend Model used in the main analyses.

[34] The positive effect of TIMESREP continues to hold when measuring stickiness on KAM level rather than for the entire auditor's report. However, data on stickiness of KAMs is only available from the second post-ISA 700 period onwards, significantly reducing the number of observations (results are untabulated).

[35] In line with prior studies on textual characteristics of KAM disclosures (e.g. Seebeck & Kaya, 2022; Smith, in press) the adjusted *R*-squared varies between 0.225 and 0.385, suggesting that a reasonable percentage of the variance in similarity scores is explained by the input variables.

[36] Given that 98.2% of the firms in the sample are audited by the Big 4 auditors, this finding should be interpreted with caution. For instance, it cannot be concluded from the results that the auditor's reports of Big 4 audit firms are of lower quality.

[37] I note that a general limitation of the Cosine similarity approach is that different yet connotatively similar words are not considered to be similar (e.g. 'stocks' and 'shares').

[38] The Cosine similarity approach is one of the most common uses of the Vector Space Model (VSM) described by Salton et al. (1975). VSM is an algebraic model for representing text documents as vectors. It was developed to compare strings of text or documents and has a wide range of applications in Natural Language Processing.

[39] For more details on the approach and its applicability to financial reports, including examples and validation, see Brown and Tucker (2011).

[40] The TF weighting addresses the concern that the similarity score may remain unaffected if an auditor merely increases boilerplate disclosure. This is because boilerplate language typically consists of common words that receive only very little weight in the TF-IDF approach.

[41] Brown and Tucker (2011) argue that the TF-IDF weighting eliminates the need for removing stop words because most stop words are common words which receive only very little weight. However, in additional analyses I find that some stop words are rarely used in auditor's reports (e.g. 'against' and 'between').

REFERENCES

Asbahr, K., & Ruhnke, K. (2019). Real effects of reporting key audit matters on auditors' judgment and choice of action. *International Journal of Auditing*, *23*(2), 165–180. https://doi.org/10.1111/ijau.12154

Auditing and Assurance Standards Board (AASB). (2017). CAS 701, communicating key audit matters in the independent auditor's report. https://www.cpacanada.ca/en/business-and-accountingresources/audit-and-assurance/canadian-auditing-standards-cas/publications/audit-and-assurance-alertcas-701

Bédard, J., Coram, P., Espahbodi, R., & Mock, T. J. (2016). Does recent academic research support changes to audit reporting standards? *Accounting Horizons*, *30*(2), 255–275. https://doi.org/10.2308/aach51397

Bédard, J., Gonthier-Besacier, N., & Schatt, A. (2014). Costs and benefits of reporting key audit matters in the audit report: The French experience. *International Symposium on Audit Research*, *24*, 1–24.

Bernard, D., Burgstahler, D., & Kaya, D. (2018). Size management by European private firms to minimize proprietary costs of disclosure. *Journal of Accounting and Economics*, *66*(1), 94–122. https://doi.org/10.1016/j.jacceco.2018.03.001

Blankespoor, E., deHaan, E., & Marinovic, I. (2020). Disclosure processing costs, investors' information choice, and equity market outcomes: A review. *Journal of Accounting and Economics*, *70*(2–3), 101344. https://doi.org/10.1016/j.jacceco.2020.101344

Brasel, K., Doxey, M. M., Grenier, J. H., & Reffett, A. (2016). Risk disclosure preceding negative outcomes: The effects of reporting critical audit matters on judgments of auditor liability. *The Accounting Review*, *91*(5), 1345–1362. https://doi.org/10.2308/accr-51380

Brown, S. V., & Knechel, R. (2016). Auditor–client compatibility and audit firm selection. *Journal of Accounting Research*, *54*(3), 725–775. https://doi.org/10.1111/1475-679X.12105

Brown, S. V., & Tucker, J. W. (2011). Large-sample evidence on firms' year-over-year MD&A modifications. *Journal of Accounting Research*, *49*(2), 309–346. https://doi.org/10.1111/j.1475-679X.2010.00396.x

Burke, J. J., Hoitash, R., Hoitash, U., & Xiao, S. X. (2023). The disclosure and consequences of US critical audit matters. *The Accounting Review*, *98*(2), 59–95. https://doi.org/10.2308/TAR-2021-0013

Bushee, B. J., & Leuz, C. (2005). Economic consequences of SEC disclosure regulation: Evidence from the OTC bulletin board. *Journal of Accounting and Economics*, *39*(2), 233–264. https://doi.org/10.1016/j.jacceco.2004.04.002

Campbell, J. L., Chen, H., Dhaliwal, D. S., Lu, H. M., & Steele, L. B. (2014). The information content of mandatory risk factor disclosures in corporate filings. *Review of Accounting Studies*, *19*(1), 396–455. https://doi.org/10.1007/s11142-013-9258-3

Carlé, T., Pappert, N., & Quick, R. (2023). Text similarity, boilerplates and their determinants in key audit matters disclosure.*Corporate Ownership & Control*, *20*(2), 49–62. https://doi.org/10.22495/cocv20i2art4

Cazier, R. A., McMullin, J. L., & Treu, J. S. (2021). Are lengthy and boilerplate risk factor disclosures inadequate? An examination of judicial and regulatory assessments of risk factor language. *The Accounting Review*, *96*(4), 131–155. https://doi.org/10.2308/TAR-2018-0657

Chinese Institute of Certified Public Accountants (CICPA). (2016). Chinese standard on auditing 1504: Communicating key audit matters in the independent auditor's report. https://www.cicpa.org.cn/xxfb/news/201612/W020210118390501351623.pdf

Christensen, B. E., Glover, S. M., & Wolfe, C. J. (2014). Do critical audit matter paragraphs in the audit report change nonprofessional

investors' decision to invest? *Auditing: A Journal of Practice & Theory*, *33*(4), 71–93. https://doi.org/10.2308/ajpt-50793

Christensen, B. E., Neuman, S. S., & Rice, S. C. (2019). The loss of information associated with binary audit reports: Evidence from auditors' internal control and going concern opinions. *Contemporary Accounting Research*, *36*(3), 1461–1500. https://doi.org/10.1111/1911-3846.12470

Deneuve, E., Filip, A., & Jeny, A. (2023). Detecting informative value in key audit matters: the importance of dissimilar kam risk descriptions. *Workingpaper*, Retrieved November 24, 2023 from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4643764.

DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, *48*, 147–160. https://doi.org/10.2307/2095101

Dyer, T., Lang, M., & Stice-Lawrence, L. (2017). The evolution of 10-K textual disclosure: Evidence from latent Dirichlet allocation. *Journal of Accounting and Economics*, *64*(2–3), 221–245. https://doi.org/10.1016/j.jacceco.2017.07.002

Dyer, T. A., Roulstone, D. T., & van Buskirk, A. (2023). *Disclosure similarity and future stock return comovement*. Management Science. https://doi.org/10.1287/mnsc.2023.4915

Elliott, W. B., Fanning, K., & Peecher, M. E. (2020). Do investors value higher financial reporting quality, and can expanded audit reports unlock this value? *The Accounting Review*, *95*(2), 141–165. https://doi.org/10.2308/accr-52508

Financial Reporting Council (FRC). (2009). *Auditing practices board's ethical standard 3 (revised)*. FRC. https://www.frc.org.uk/getattachment/2cc10083-8576-48a5-9b9a-50c9f1c41441/ES-3-(Revised)-Long-Association-with-the-Audit-Eng.pdf

Financial Reporting Council (FRC). (2013a). *International standard on auditing (UK and Ireland) 700: The independent auditor's report on financial statements (June 2013)*. FRC. https://www.frc.org.uk/getattachment/501de004-b616-43c3-8d65-aeaebde19f8d/ISA-700-(UK-and-Ireland)-700-(Revised)-Independent-auditors-report-June-2013.pdf

Financial Reporting Council (FRC). (2013b). *Feedback statement. Feedback statement on consultation on proposed revision to ISA (UK and Ireland) 700*. FRC. https://www.frc.org.uk/getattachment/95da69ad-6996-496b-adb7-b0f1522badc9/Feedback-Statement-for-publication.pdf

Financial Reporting Council (FRC). (2015). *Extended auditor's reports—A review of experience in the first year*. FRC. http://www.cafr.ro/uploads/3.%20Extended-auditors-reports-0773.pdf

Financial Reporting Council (FRC). (2016a). *International standard on auditing (UK) 700: Forming an opinion and reporting on financial statements (June 2016)*. FRC. https://www.frc.org.uk/getattachment/a08b0906-f40c-4735-bbc7-45908bee2b32/ISA-(UK)-700_Revised-June-2016.pdf

Financial Reporting Council (FRC). (2016b). *International standard on auditing (UK) 701: Communicating key audit matters in the independent auditor's report (June 2016)*. FRC. https://www.frc.org.uk/getattachment/b250cf61-407f-4b1e-9f1c-e959174e1426/ISA-(UK)-701.pdf

Financial Reporting Council (FRC). (2020). *Bulletin: Guidance for auditors and matters considered where engagements are affected by coronavirus (COVID-19)*. FRC. https://www.frc.org.uk/getattachment/ae0044e3-a7bf-4b75-8aa2-4e39e20f525b/Bulletin-Coronavirus-Guidance-December-2020.pdf

Financial Reporting Council (FRC). (2022). *International standard on auditing (UK) 701 (revised November 2019) (updated May 2022): Communicating key audit matters in the independent auditor's report*. FRC. https://www.frc.org.uk/document-library/audit-and-assurance/2020/isa-(uk)-701-updated-january-2020

Francis, J. R. (2019). What do we learn by studying accounting and governance practices in other countries? Discussion of "impact of auditor report changes on financial reporting quality and audit costs: Evidence from the United Kingdom". *Contemporary Accounting Research*, *36*(3), 1531–1539.

FRC (Financial Reporting Council). (2016c). *Extended auditor's reports – A further review of experience*. FRC. http://www.cafr.ro/uploads/3.%20Extended-auditors-reports-0773.pdf

Gimbar, C., Hansen, B., & Ozlanski, M. E. (2016). The effects of critical audit matter paragraphs and accounting standard precision on auditor liability. *The Accounting Review*, *91*(6), 1629–1646. https://doi.org/10.2308/accr-51382

Gray, G. L., Turner, J. L., Coram, P. J., & Mock, T. J. (2011). Perceptions and misperceptions regarding the unqualified auditor's report by financial statement preparers, users, and auditors. *Accounting Horizons*, *25*(4), 659–684. https://doi.org/10.2308/acch-50060

Gutierrez, E., Minutti-Meza, M., Tatum, K. W., & Vulcheva, M. (2018). Consequences of adopting an expanded auditor's report in the United Kingdom. *Review of Accounting Studies*, *23*(4), 1543–1587. https://doi.org/10.1007/s11142-018-9464-0

Hanley, K., & Hoberg, G. (2010). The information content of IPO prospectuses. *The Review of Financial Studies*, *23*(7), 2821–2864. https://doi.org/10.1093/rfs/hhq024

Hoberg, G., & Phillips, G. (2010). Product market synergies and competition in mergers and acquisitions: A text-based analysis. *The Review of Financial Studies*, *23*, 3773–3811. https://doi.org/10.1093/rfs/hhq053

Hope, O. K., Hu, D., & Lu, H. (2016). The benefits of specific risk-factor disclosures. *Review of Accounting Studies*, *21*(4), 1005–1045. https://doi.org/10.1007/s11142-016-9371-1

Huang, A. H., Zang, A. Y., & Zheng, R. (2014). Evidence on the information content of text in analyst reports. *The Accounting Review*, *89*(6), 2151–2180. https://doi.org/10.2308/accr-50833

Institute of Chartered Accountants in England and Wales (ICAEW). (2017). *The start of a conversation—The extended auditor report*. ICAEW. https://www.icaew.com/technical/audit-and-assurance/faculty/the-extended-audit-report

International Auditing and Assurance Standards Board (IAASB). (2015). *International standard on auditing 700 (revised), forming an opinion and reporting on financial statements*. IAASB. https://www.ifac.org/system/files/publications/files/ISA-700-Revised_8.pdf

International Auditing and Assurance Standards Board (IAASB). (2020). *Auditor reporting in the current evolving environment due to COVID-19*. IAASB. https://www.ifac.org/system/files/publications/files/Staff-Alert-Auditor-Reporting-Final.pdf

Irish Auditing and Accounting Supervisory Authority (IAASE). (2020). Statutory reporting quality. *ISA 701 – Thematic Review*. IAASE. https://www.iaasa.ie/News/2020/IAASA-has-today-published-its-ISA-701-Thematic-Rev

Kaya, D., & Koch, M. (2015). Countries' adoption of the international financial reporting standard for small and medium-sized entities (IFRS for SMEs)—Early empirical evidence. *Accounting and Business Research*, *45*(1), 93–120. https://doi.org/10.1080/00014788.2014.969188

Kelton, A. S., & Montague, N. R. (2018). The unintended consequences of uncertainty disclosures made by auditors and managers on nonprofessional investor judgments. *Accounting, Organizations and Society*, *65*, 44–55. https://doi.org/10.1016/j.aos.2017.10.001

Kravet, T., & Muslu, V. (2013). Textual risk disclosures and investors' risk perceptions. *Review of Accounting Studies*, *18*(4), 1088–1122. https://doi.org/10.1007/s11142-013-9228-9

Lang, M., & Stice-Lawrence, L. (2015). Textual analysis and international financial reporting: Large sample evidence. *Journal of Accounting and Economics*, *60*(2–3), 110–135. https://doi.org/10.1016/j.jacceco.2015.09.002

Lennox, C. S., Schmidt, J. J., & Thompson, A. M. (2023). Why are expanded audit reports not informative to investors? Evidence from the United Kingdom. *Review of Accounting Studies*, 28(2), 497–532.

Linck, J. S., Netter, J. M., & Yang, T. (2009). The effects and unintended consequences of the Sarbanes-Oxley Act on the supply and demand for directors. *The Review of Financial Studies*, *22*(8), 3287–3328. https://doi.org/10.1093/rfs/hhn084

Loughran, T., & McDonald, B. (2014). Measuring readability in financial disclosures. *The Journal of Finance*, *69*(4), 1643–1671. https://doi.org/10.1111/jofi.12162

Loughran, T., & McDonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, *54*(4), 1187–1230. https://doi.org/10.1111/1475-679X.12123

Loughran, T., & McDonald, B. (2020). Textual analysis in finance. *Annual Review of Financial Economics*, *12*, 357–375. https://doi.org/10.1146/annurev-financial-012820-032249

Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology*, *83*(2), 340–363. https://doi.org/10.1086/226550

Miller, G. S. (2017). Discussion of "the evolution of 10-K textual disclosure: Evidence from latent Dirichlet allocation". *Journal of Accounting and Economics*, *64*(2–3), 246–252. https://doi.org/10.1016/j.jacceco.2017.07.004

Minutti-Meza, M. (2021). The art of conversation: The expanded audit report. *Accounting and Business Research*, *51*(5), 548–581. https://doi.org/10.1080/00014788.2021.1932264

Moroney, R., Phang, S. Y., & Xiao, X. (2021). When do investors value key audit matters? *The European Accounting Review*, *30*(1), 63–82. https://doi.org/10.1080/09638180.2020.1733040

Norris, F. (2014). Holding auditors accountable on reports. *The New York Times*. May 9, 2014, Section B, p. 1. https://www.nytimes.com/2014/05/09/business/holding-auditors-accountable.html

Parsons, T. (1956). Suggestions for a sociological approach to the theory of organizations-I. *Administrative Science Quarterly*, *1*, 63–85. https://doi.org/10.2307/2390840

Parsons, T. (1960). Pattern variables revisited: A response to Robert Dubin. *American Sociological Review*, *25*, 467–483. https://doi.org/10.2307/2092932

Pelzer, J. R. E. (2021). Processing change: A qualitative study examining the frontstage and backstage of audit firms contemplating the implementation of critical audit matters. *International Journal of Auditing*, *25*(3), 769–796. https://doi.org/10.1111/ijau.12248

Pinto, I., Morais, A. I., & Quick, R. (2020). The impact of the precision of accounting standards on the expanded auditor's report in the European Union. *Journal of International Accounting, Auditing and Taxation*, *40*, 100333. https://doi.org/10.1016/j.intaccaudtax.2020.100333

Porumb, V., Zengin-Karaibrahimoglu, Y., Lobo, G. J., Hooghiemstra, R., & de Waard, D. (2021). Expanded auditor's report disclosures and loan contracting. *Contemporary Accounting Research*, *38*(4), 3214–3253. https://doi.org/10.1111/1911-3846.12697

Public Company Accounting Oversight Board (PCAOB). (2017). *The auditor's report on an audit of financial statements when the auditor expresses an unqualified opinion (AS 3101), PCAOB. Release no. 2017–001*. PCAOB.

Reid, L. C., Carcello, J. V., Li, C., & Neal, T. L. (2019). Impact of auditor report changes on financial reporting quality and audit costs: Evidence from the United Kingdom. *Contemporary Accounting Research*, *36*(3), 1501–1539. https://doi.org/10.1111/1911-3846.12486

Reilly, K. (2014). *Panel: Critical audit matters related to the audits of large companies. PCAOB auditor's reporting model public meeting on April 2, 2014*. SEC. Retrieved September 2, 2022 from: https://pcaob-assets.azureedge.net/pcaob-dev/docs/default-source/rulemaking/docket034/04022014_transcript.pdf?sfvrsn=a966a842_0

Salton, G., Wong, A., & Yang, C. S. (1975). A vector space model for automatic indexing. *Communications of the ACM*, *18*(11), 613–620. https://doi.org/10.1145/361219.361220

Securities and Exchange Commission (SEC). (1998). *Plain English disclosure, 63 FR 6370*. SEC. https://www.govinfo.gov/app/details/FR-1998-02-06/98-2889

Securities and Exchange Commission (SEC). (2005). *Securities and exchange commission final rule*, release no. 33–8591 *1–468*. SEC. https://www.sec.gov/rules/final/33-8591.pdf

Securities and Exchange Commission (SEC). (2016). *Business and financial disclosure required by regulation S-K*, release no. 33–10064; *34–77599*; *File No. S7-06-16*, *1–341*. SEC. https://www.sec.gov/rules/concept/2016/33-10064.pdf

Seebeck, A., & Kaya, D. (2022). The power of words: An empirical analysis of the communicative value of extended auditor reports. *The European Accounting Review*, *35*(5), 1185–1215.

Seebeck, A., & Vetter, J. (2022). Not just a gender numbers game: How board gender diversity affects corporate risk disclosure. *Journal of Business Ethics*, *177*(2), 395–420. https://doi.org/10.1007/s10551-020-04690-3

Sirois, L.-P., Bédard, J., & Bera., P. (2018). The informational value of key audit matters in the auditor's report: Evidence from an eye-tracking study. *Accounting Horizons*, *32*(2), 141–162. https://doi.org/10.2308/acch-52047

Smith, K. W. (2023). Tell me more: A content analysis of expanded auditor reporting in the United Kingdom. *Accounting, Organizations and Society*, *108*, 101456. https://doi.org/10.1016/j.aos.2023.101456

Zeng, Y., Zhang, J. H., Zhang, J., & Zhang, M. (2021). Key audit matters reports in China: Their descriptions and implications of audit quality. *Accounting Horizons*, *35*(2), 167–192. https://doi.org/10.2308/HORIZONS-19-189

AUTHOR BIOGRAPHY

Andreas Seebeck is an Assistant Professor at the Constructor University in Bremen. His major research interests include empirical financial accounting, auditing, and capital market research.

How to cite this article: Seebeck, A. (2024). Back to where it started?—Do expanded auditor's reports become sticky, generic and boilerplate over time? *International Journal of Auditing*, *28*(3), 536–561. https://doi.org/10.1111/ijau.12343

APPENDIX A: VARIABLE DEFINITIONS

Variable	Definition
Dependent variables	
Year-over-year similarity of auditor's reports	
STICKINESS	The year-over-year Cosine similarity score for the full auditor's report. In the main specification, the TF-IDF weighting scheme is applied and stop words are excluded.
Within-peer group similarity of KAM descriptions	
SIM_YEAR	The mean Cosine similarity score of a client's KAM descriptions to all other KAM descriptions issued in a given post-ISA 700 reporting period (percentile ranking; 1 to 100). In the main specification, the TF-IDF weighting scheme is applied and stop words are excluded.
SIM_AUDITOR	The mean Cosine similarity score of a client's KAM descriptions to all other KAM descriptions issued by the same audit firm in a given post-ISA 700 reporting period (percentile ranking; 1 to 100). In the main specification, the TF-IDF weighting scheme is applied and stop words are excluded.
SIM_RISK	The mean Cosine similarity score of a client's KAM descriptions to all other KAM descriptions that address the same risk type in a given post-ISA 700 reporting period (percentile ranking; 1 to 100). In the main specification, the TF-IDF weighting scheme is applied and stop words are excluded.
SIM_PEER	Aggregate within-peer group similarity measure, calculated as the average of the percentiles of SIM_YEAR, SIM_AUDITOR and SIM_RISK (1 to 100).
Boilerplate language in KAM sections	
BOILERPLATE	The number of boilerplate phrases in a KAM section. Boilerplate phrases are defined as four-grams (five-grams, six-grams) that occur in more than 50% of the sample firms' KAM sections. In the main specification, stop words are excluded.
Aggregate similarity measures	
SIM_ALL	Aggregate similarity measure, calculated as the average of the percentiles of STICKINESS, SIM_PEER and BOILERPLATE (1 to 100).
ABN_SIM_ALL	Aggregate similarity measure that measures the abnormal similarity contained in SIM_ALL. The expected similarity for a given auditor's report is determined based on Equation (1). The abnormal similarity is the difference between the observed similarity values and the estimated values.
Equity market effects	
ABVOL	The natural logarithm of the 3-day abnormal trading volume around the report filing date scaled by shares outstanding. Estimation period trading volume is measured over a 40 days trading period beginning 61 days before the earnings announcement date, following Gutierrez et al. (2018).
|CAR|	The absolute value of the sum of the 3-day abnormal returns around the report filing date. Company returns are calculated using the formula: (Price Closet – Price Closet-1)/Price Closet-1 minus value-weighted same-day returns for prime listed firms on the London Stock Exchange (LSE), following Gutierrez et al. (2018).
Test variable	
TIMESREP	The number of times an expanded auditor's report has been reported under new ISA 700 requirements for a given client. Categorical variable with the possible positive integer values of 1 to 6.
Control variables	
BIGN	An indicator variable that is equal to one if the auditor is a Big 4 audit firm, and zero otherwise.
BUSY	An indicator variable that is equal to one if the financial statement date is between December and March, and zero otherwise.
CFO	The natural logarithm of a firm's operating cash flow.
CHAUD	An indicator variable that is equal to one if the client changed its auditor (i.e. audit firm) compared to the previous year, and zero otherwise.
CHNI	Difference in net income before extraordinary items between year t and year $t-1$, divided by total assets in year $t-1$.
DELAY	The number of calendar days between a firm's fiscal year-end and the date of the auditor's report.
DOWNSIZE	An indicator variable that is equal to one if total assets decrease by at least one-tenth from the previous year, and zero otherwise.
EXPERT	An indicator variable that is equal to one if the auditor is an industry expert, and zero otherwise. An auditor is deemed an industry expert if at least 50% of all firms in the sample that belong to a specific industry are audited by the auditor. I only consider those industries for which I have at least four firm-year observations.
INCORP	Number of years since incorporation.
LENGTH	The natural logarithm of the total number of words in the KAM section.
LEV	The long-term debt divided by total equity at fiscal year-end.

(Continues)

Variable	Definition
LOSS	An indicator variable that is equal to one if net income is negative, and zero otherwise.
MB	The market value divided by book value at fiscal year-end.
MOWN	An indicator variable that is equal to one if the majority of shares is held by top 10 investors, and zero otherwise.
REPORT_DISCL	The natural logarithm of the total number of words in the KAM section.
ROA	The net income prior to financing costs divided by total assets at fiscal year-end.
SIZE	The natural logarithm of total assets at the end of the fiscal year.
SPECIFICITY	The mean of seven indicator variables, which indicate the occurrence of specific words (names of persons, locations, organisations; quantitative values in percentages, money values, times and dates) in the KAM description. Each variable is equal to one if the KAM description contains specific words from the respective dimension, and zero otherwise. Specific words are identified using the Stanford NER program, following Seebeck and Kaya (2022).
STDRET	Standard deviation of daily stock returns in year *t*.
USLIST	An indicator variable that is equal to one if the firm is cross-listed on the US stock exchange, and zero otherwise.

APPENDIX B: EXAMPLES FOR SIMILARITY MEASURES

TABLE B1 Boilerplate phrases and sticky disclosure.

Auditor: William Touche, Deloitte LLP **Audit client: Sky plc (2015)**	**Auditor: Paul Franek, Deloitte LLP** **Audit client: Sky plc (2016)**
Capital expenditure *In the UK, spending on capital projects is material and the assessment and timing of whether assets meet the capitalisation criteria set out in IAS 16 Property, Plant and Equipment and IAS 38 Intangible Assets requires judgement as set out in the Group's critical accounting policies on page 93.* *In addition, determining whether there is any indication of impairment of the carrying value of assets being developed also requires judgement.* *As a result, there is a risk that expenditure on intangible and tangible non-current assets in the UK business is inappropriately capitalized against relevant accounting guidance and that assets not yet in use are not recoverable at their carrying value.*	*Capital expenditure* *The Group's **spending on capital projects is material**, as shown by the total value of additions in notes 13 and 14. **The assessment and timing of whether assets meet the capitalisation criteria set out in IAS 16 Property, Plant and Equipment and IAS 38 Intangible Assets requires judgement, as set out in the Group's critical accounting policies on page** 86.* ***In addition, determining whether there is any indication of impairment of the carrying value of assets being developed** or replaced **also requires judgement** in assessing performance against the investment business case.* ***As a result, there is a risk that** the Group's **expenditure on intangible and tangible non-current assets is inappropriately** capitalised **against relevant accounting guidance, that assets not yet in use are not recoverable at their carrying value** and that the value of existing assets made obsolete by current year additions may be impaired.*

Notes: Table B1 shows examples of the use of boilerplate phrases and sticky disclosure in KAM descriptions for Sky plc in the years 2015 and 2016. The first column shows a KAM description for Sky plc (2015). All boilerplate phrases, which are identified based on the four-gram identification strategy described in Section 4.2., are underlined.
The second column shows the description of the same KAM for Sky plc for the subsequent year (2016). The reporting in the second column is compared to the reporting in the first column. Sticky disclosure is highlighted by bolded text. Note that the word '*capitalised*' (GB-en) used in 2016 is not recognised as being similar to the word '*capitalized*' (US-en) used in 2015.

TABLE B2 Within peer group similarity.

Auditor: Mark Mullins, Deloitte LLP Audit client: Spirax–Sarco Engineering plc (2015)	Auditor: Alexander Butterworth, Deloitte LLP Audit client: Halma plc (2015)
Defined benefit pension plan assumptions	*Defined benefit pension plan assumptions*
At 31 December 2015 the net retirement benefit liability recognised in the Statement of financial position was £74 million (31 December 2014: £76 million).	*At 28 March 2015 the net retirement benefit liability recognised in the Consolidated Balance Sheet was £67 m (2014: £37 m).*
There is a risk relating to the judgements made by management in valuing the defined benefit pension liabilities including the use of key model input assumptions including the discount rate, mortality assumption and inflation level.	*There is a risk relating to judgments made by management in valuing the defined benefit pension plans including the use of key model input assumptions such as discount rates, mortality assumptions and inflation levels.*
These variables can have a material impact in calculating the quantum of the retirement benefit liability.	*These variables can have a material impact in calculating the quantum of the retirement benefit liability.*
Refer to Note 1 for the Group's policy on defined benefit plans and Note 24 for the financial disclosure including the key assumptions used in the defined benefit pension plan valuation.	*Management utilise the services of third party actuarial advisers to determine their key assumptions.*
	Details of the defined benefit pension plans are disclosed in note 28 to the accounts.
	The Audit Committee has included their assessment of this risk on page 69 and it is included in the key accounting estimates and judgments on pages 106 and 107.

Notes: Table B2 shows a comparison of two KAM descriptions for the audit clients Spirax–Sarco Engineering plc and Halma plc from the financial year 2015.

Words and phrases that are similarly used across both KAMs are underlined in the second column. The similarity score of the two KAMs is determined using Cosine similarity scores. The similarity score is considered for all three SIM_PEER measures for the given example because (1) both KAMs were reported for the same financial year, that is, 2015 (SIM-YEAR), (2) both KAMs describe the same type of risk, that is, defined benefit pension assumptions (SIM_RISK) and (3) both audit engagement partners worked for the same audit firm, that is, Deloitte LLP (SIM_AUDITOR).

APPENDIX C: COSINE SIMILARITY APPROACH AND TECHNICAL SPECIFICATIONS

This paper uses the Cosine similarity approach to measure year-over-year similarity (stickiness) and within-peer group similarity of expanded auditor reporting. According to the Cosine similarity approach, the similarity of the two reports is described by the angle between two vectors epitomising the auditor's reports:

$$\vartheta_1 = (\omega_1, \omega_2, ..., \omega_{n-1}, \omega_n) \text{ and } \vartheta_2 = (\psi_1, \psi_2, ..., \psi_{n-1}, \psi_n).$$

where ω_i and ψ_i are the frequencies of the words $i \in [1,n]$ in document 1 (ϑ_1) and document 2 (ϑ_2), respectively.[37]

The formula to compute the Cosine similarity score is:

$$\textit{Cosine similarity} = \cos(\theta) = \frac{\vartheta_1}{\|\vartheta_1\|} \cdot \frac{\vartheta_2}{\|\vartheta_2\|} = \frac{\vartheta_1 \cdot \vartheta_2}{\|\vartheta_1\| \, \|\vartheta_2\|}$$

where θ is the angle between ϑ_1 and ϑ_2, $(\cdot)$ is the dot product operator, and $\|\vartheta_1\|$, $\|\vartheta_2\|$ are the vector lengths of ϑ_1 and ϑ_2, respectively.[38] The similarity score is a value between 0 (no similarity) and 1 (complete similarity).[39]

In the main analysis, I use the 'term frequency-inverse document frequency' (TF-IDF) weighting scheme to assign different weights to the words used in the corpora which reflect the importance of a word to the document. The TF-IDF similarity score is calculated as follows:

$$\textit{Cosine similarity}_{TF-IDF} = \textit{Cosine similarity} * \ln(M/m)$$

where M is the number of total reports in the sample and m is the number of reports where the respective word appears. The TF-IDF value increases proportionally to the number of times a word appears in the document and is offset by the number of documents in the corpus which contain the word in question. Thus, with the TF-IDF concept, the value of Cosine similarity depends on the rest of the sample, with common words being weighted less. Vice versa, words that are unique to a report are heavily weighted and are therefore considered more informative. For instance, the term 'misstatement' occurs in almost all documents, meaning it conveys only limited information. By contrast, the word 'liquidation' is not commonly used and is thus very indicative of the content of the few documents where it occurs. In additional analyses, I use the 'term frequency' (TF) weighting scheme.[40] Moreover, Cosine similarity can be measured with or without considering stop words, that is, words such as articles and prepositions. which are necessary for writing a text with correct grammar. Methods which both exclude (e.g. Brown & Tucker, 2011) and include stop words (e.g. Hanley & Hoberg, 2010) are frequently used in prior research.[41] In the main analysis, I exclude stop words. In additional analyses, I combine both approaches (i.e. including and excluding stop words) with the different weighting schemes described above (i.e. TF-IDF and TF), resulting in four different *STICKINESS* scores as indicated by the suffixes in Table 5A. My results remain qualitatively the same for all specifications.