

It's OK to pay well, if you write well: The effects of remuneration disclosure readability

Hemmings, Danial; Hodgkinson, Lynn; Williams, Gwion

Journal of Business Finance and Accounting

DOI:
[10.1111/jbfa.12431](https://doi.org/10.1111/jbfa.12431)

Published: 01/05/2020

Peer reviewed version

[Cyswllt i'r cyhoeddiad / Link to publication](#)

Dyfyniad o'r fersiwn a gyhoeddwyd / Citation for published version (APA):
Hemmings, D., Hodgkinson, L., & Williams, G. (2020). It's OK to pay well, if you write well: The effects of remuneration disclosure readability. *Journal of Business Finance and Accounting*, 47(5-6), 547-586. <https://doi.org/10.1111/jbfa.12431>

Hawliau Cyffredinol / General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

**It's OK to pay well, if you write well:
The effects of remuneration disclosure readability**

Danial Hemmings^{a,†}, Lynn Hodgkinson^a, Gwion Williams^a

^a *Bangor University, Bangor Business School, Hen Goleg, College Road,
Bangor, Gwynedd, United Kingdom, LL57 2DG.*

[†] Corresponding author. Email address: d.hemmings@bangor.ac.uk

This version: 6th December 2019

Acknowledgements

Comments, suggestions, and encouragement from Ozlem Arikan, Georgios Voulgaris, and Lucrezia Fattobene, and seminar participants at Warwick Business School, Newcastle University Business School, Strathclyde Business School, Bangor Business School, the 2016 Gregynog A&F colloquium, the 2016 Wolpertinger Conference, the 2017 BAFA Annual Conference, and the BAFA Northern Area Group Annual Conference, and Australian seminar participants are gratefully acknowledged.

Data Availability Statement

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

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

**It's OK to pay well, if you write well:
The effects of remuneration disclosure readability**

Abstract

We examine whether, and how, shareholders' votes in the Say-on-Pay (SOP) are affected by the readability of the Compensation Discussion and Analysis (CD&A). Despite the SEC's Plain English requirement, qualitative disclosures on executive remuneration are generally long and complex. Extant evidence on whether low readability results in higher or lower shareholder dissent in the SOP however is ambiguous. We resolve this debate by demonstrating that the effects of readability on SOP voting are heterogeneous; while obfuscation may reduce dissent when CEO compensation is close to 'normal' levels, diminished readability results in increased scepticism when pay levels are clearly excessive. The moderating role of readability is most pronounced for firms with less sophisticated shareholders, consistent with readability acting as a heuristic cue. Our results are robust to propensity score matching, and are less pronounced (1) when shareholders have less time to review the CD&A, and (2) when shareholders are distracted by competing AGMs, suggesting they are driven by readability, directly. Overall, our results highlight that greater use of Plain English in remuneration disclosures can have a substantial persuasive impact on shareholders.

Keywords: Executive compensation, Say-on-Pay, Compensation Discussion and Analysis, Plain English, Readability.

JEL Classification: G34, G38, J33, M41, M48

1. Introduction

Despite the SEC's requirement that firms produce a Plain English report on executive remuneration, the typical Compensation Discussion and Analysis (CD&A) remains complex and un-engaging (Larcker et al., 2015). Accessible and informative disclosure in the CD&A is important as it serves as a basis for the Say-on-Pay (SOP) shareholder vote on executive remuneration. The SOP provides shareholders a voice on whether they approve the compensation of Named Executive Officers.¹ With the aim of fostering shareholder engagement, the SEC encourages firms to produce a CD&A that clearly communicates the firm's compensation story, and which "*put[s] into perspective the numbers and narrative that follow it*" (SEC, 2006, p. 29). Counter to this idea, however, Larcker et al.'s (2015) survey of institutional investors finds that only 38% consider the typical compensation disclosure to be clear and easy to understand.

While very few prior studies in Accounting and Finance have examined the text of remuneration disclosures (El-Haj et al., 2019b), extant evidence suggests that poor readability has ambiguous effects on shareholder voting patterns in the SOP. For example, Hooghiemstra et al. (2017) find that diminished readability of the Directors Remuneration Report (DRR) can successfully obfuscate problematic CEO pay in the UK, thereby leading to lower SOP dissent, on average, in the face of excessive CEO pay. On the other hand, they find that low readability leads to increased dissent at firms with high institutional ownership, suggesting that low readability may, alternatively, increase scepticism regarding the appropriateness of pay arrangements in some cases. Lo et al. (2014) report evidence of a negative relationship between CD&A readability and SOP dissent, on average, for S&P 1500 firms, suggesting higher overall levels of scepticism accompany low readability in the US. Conversely, Balsam et al. (2016),

¹ Named Executive Officers comprise the Chief Executive Officer, Chief Financial Officer and the next three most highly paid executives.

also using US data, do not find a significant relationship between CD&A readability and SOP dissent, overall.

Our study contributes to this debate by demonstrating that the effect of readability on SOP dissent is contingent on the perceived degree of excessiveness in CEO pay. Intuitively, our results indicate that low readability leads shareholders to become increasingly sceptical of apparent CEO pay excesses, as the latter increases. The obfuscation effect of diminished readability in reducing SOP dissent, identified by Hooghiemstra et al. (2017), appears limited to cases where pay excesses are moderate, but gives-way to increased scepticism and higher SOP dissent when CEO pay is prima-facie more clearly excessive.

The manner by which disclosure readability affects investors' decision making has received increased attention in recent years. For example, it has been shown that readability affects the valuation judgments of equity-holders (Rennekamp, 2012; Hsieh et al., 2016; Hwang and Kim, 2017), as well as the risk perceptions of debt-holders (Bosnall and Miller, 2017). The potential impact of corporate disclosure readability on shareholder monitoring and voting decisions (i.e., shareholder 'voice'), however, remains largely unexplored. Arrangements for executive pay is a setting where there are incentives for managers to foster support from shareholders through strategic communication or symbolic action. For example, evidence from "*just vote no*" campaigns that preceded SOP suggests that mere public displays of shareholder dissatisfaction with pay arrangements can incentivise directors to take swift action to avoid embarrassment (Del Guercio et al., 2008), in some cases even ousting the reigning CEO (Ertimur et al., 2011). While the SOP in the U.S. is currently non-binding, high SOP dissent can lead to public shaming (Cai and Walkling, 2011) and enhanced external scrutiny (Ertimur et al., 2013). The SOP can therefore provide a catalyst for enhanced shareholder outreach and engagement.

It is generally accepted that low readability increases the costs to shareholders of collecting and processing information (Bloomfield, 2002). Since investors' attention is a scarce resource (Hirshleifer and Teoh, 2003), allocation of time and effort to processing information from corporate documents may be limited when these costs outweigh the benefits of being better informed. In the context of the SOP, retail investors,² in particular, may consider that their vote is unlikely to be pivotal, and thus view only slight potential benefit from a thorough review of compensation arrangements. Accordingly, small shareholders appear disinclined to engage meaningfully with the SOP unless pay arrangements are clearly inappropriate, and tend not to vote 'against' in the SOP unless executive pay is highly excessive (Stathopoulos and Voulgaris, 2016). It is reasonable, therefore, to expect that more complex CD&A may 'put off' shareholders from challenging executive pay arrangements, unless they prima-facie have a strong reason to believe that pay arrangements are problematic.

Dissuading investors from extracting information may be desirable to managers when there is information that they would prefer to keep hidden (Curtis, 1998). From the preparers' perspective, managers may therefore engage in obfuscation, for example by lowering disclosure readability, when there are incentives to conceal information from shareholders or other stakeholders, such as labour unions (Frantz et al., 2013). Managers' intent may even be "*to leave readers confused and to put them off probing further*" (Merkl-Davies and Brennan, 2007, p. 12). Laksmana et al. (2012) find evidence of obfuscation when CD&A were first introduced, however they argue further that "*heightened public scrutiny and regulatory oversight*" (Laksmana et al., 2012, p. 201) resulted in the disappearance of this effect from as

² Retail investors (i.e., non-institutional investors) are likely to own only a small portion of the firm, which renders many monitoring activities cost-ineffective (see Grossman and Hart, 1980; Strickland et al., 1996). They are also likely to be less-sophisticated compared with institutional (i.e., professional) investors.

early as the second year of the CD&A, suggesting deliberate obfuscation of the CD&A to be unsuccessful.

In contrast to the obfuscation hypothesis, it is well documented in the Psychology literature that processing fluency (the ease with which information is gathered and processed) tends to positively affect perceptions of authors' credibility (Rennekamp, 2012). A difficult to read CD&A may thus intensify shareholders' suspicions that compensation arrangements are inappropriate, and that they cannot rely on the firm to act in their interest (Hwang and Kim, 2017). The processing fluency heuristic has been shown to affect investors' valuation decisions. For example, Rennekamp (2012) provides experimental evidence that investors weigh information more heavily within valuation judgments when presented in plainer English. Consistently, Hwang and Kim (2017) find that closed-end funds trade at a greater discount to their fundamental value when they produce less readable reports, and argue that the perceived value of delegated management is lower when they make disclosures that are difficult to read.

The obfuscation and processing fluency effects of low disclosure readability are related, yet contradictory effects. Both obfuscation and processing fluency effects derive from the notion that low readability makes it more difficult for readers to extract the information content from disclosures. Yet, while the obfuscation hypothesis predicts that low readability makes shareholders *less* likely to challenge management, the processing fluency hypothesis suggests that low readability increases suspicion, and therefore shareholders will be *more* likely to challenge management as a result. We attempt to disentangle these effects by arguing that readability effects are likely to be heterogeneous, and depend on whether shareholders *prima-facie* consider executive pay levels to be problematic, and have therefore a more pressing need for adequate justification. Specifically, we expect the obfuscation effect of low readability in reducing SOP dissent is more likely to dominate when CEO pay otherwise appears in-line with benchmarks. When CEO pay is clearly excessive (e.g., well above the sector), shareholders'

default position is to challenge management, unless the CD&A tells a convincing ‘compensation story’ – i.e., low readability hinders shareholders discovery of an adequate justification.

We isolate CD&A text from 2,686 proxy statements of S&P 1500 firms holding a SOP vote on compensation awarded between fiscal years 2010 and 2014, and measure readability using the Bog Index (following Bonsall et al., 2017), as well as the Gunning-Fog Index. We observe that the readability of the typical CD&A is very low; on average, *Bog* is estimated at 96.5, which indicates ‘poor’ readability according to the index; we estimate *Fog* to be 23.6 on average, which suggests that 23.6 years of formal education are required in order to be able to comprehend the typical CD&A on first reading. Our main results indicate that CD&A readability plays a significant moderating role on the link between CEO excess compensation (*Excess comp*) and SOP dissent, consistent with the processing fluency effect of low readability dominating over the obfuscation effect at high levels of *Excess comp*. Specifically, we find that the conditional incremental probability of receiving a high level of dissent (>30% votes cast ‘against’) given *Excess comp*, is significantly higher for firms with a *High Bog* or *High Fog* CD&A (greater than median *Bog* or *Fog*). The economic significance of this result appears substantial. We estimate the probability of high dissent for a firm with *Excess comp* of two standard deviations above the mean to be 27.5% when CD&A *Bog* is high, compared to 17.8% when low; and 33.9% (13.8%) when *Fog* is high (low). We also find the results are more pronounced for firms with low institutional ownership which is consistent with heuristic processing.

Given that our results may alternatively be driven by underlying agency problems resulting in both higher *Excess comp* and lower CD&A readability, we conduct propensity score matching based on the probability of poor readability and find our results to be robust. Since our hypotheses are based on the presumption that shareholders have sufficient time to thoroughly

review the CD&A, we also conduct further sensitivity tests based on the notion that shareholders' attention is often constrained (Hirshleifer & Teoh, 2003). We report that the readability effect weakens when (i) short notice is given between publication of the CD&A and the AGM, and (ii) when a high number of other AGMs occur at the same time. Since shareholders are less likely to have time to review each CD&A in these cases, this adds further support to the effect being driven by readability, directly.

To our knowledge, our study is also the first to examine whether remuneration report readability affects shareholders' participation in the SOP (as opposed to the level of dissent). This is important as it speaks directly to the underlying intention for the Plain English requirement, that is, to increase engagement by unsophisticated investors. To this end, our results highlight lower engagement by non-institutional shareholders in the SOP, and that this is unaffected by variations in CD&A readability and excess CEO compensation. We do, however, corroborate the views from Larcker et al.'s (2015) survey that CD&A are, in general, written in highly complex language, and thus conjecture that preconceptions regarding the general complexity of CD&A and compensation arrangements may be putting off some shareholders from engaging with the disclosures and the SOP altogether.

Finally, we examine the robustness of our findings when incorporating the role of media coverage, on the basis that negative media coverage of executive pay may co-occur with both SOP dissent (Hooghiemstra et al., 2015) and low CD&A readability. We find that firms with more excessive CEO compensation are more likely to attract negative media coverage, however we observe the effects of CD&A readability on SOP votes to be distinct from that of negative media reports.

Our paper contributes to growing empirical literature on the readability of compensation disclosures by demonstrating that the effects of readability on SOP voting are, in fact,

heterogeneous, and depend on the degree of pay excesses. While Hooghiemstra et al. (2017) provide evidence that poor readability reduces SOP dissent, we demonstrate that this obfuscation effect appears limited to cases where executive pay is close to ‘normal’ levels. Instead, when CEO pay is *prima facie* more clearly excessive, lower perceptions of management credibility from poor CD&A readability result in *higher* SOP dissent. Though Hooghiemstra et al. (2017) find some evidence of low readability increasing dissent for firms with particularly high levels of institutional ownership (IO), we demonstrate that, when accounting for the moderating effect of readability on *Excess comp*, this effect is also (and, in fact most) manifest for low IO firms. Our findings therefore resolve the otherwise seeming contradiction that a heuristic processing effect (i.e., processing fluency) would be exhibited by more sophisticated investors, rather than investor bases that are more likely to be unsophisticated.

Our paper is also the first to provide evidence of the processing fluency effects of readability within the context of the US SOP, where the use of Plain English in CD&A is actively encouraged by the SEC. Our findings are of practical relevance as they demonstrate that a clear and understandable compensation story is key to successfully justifying high executive pay. Conversely, from both a theoretical and regulatory perspective, our results also emphasise the possibility that firms could make strategic communication choices (i.e., enhancing the presentation of the CD&A) to inflate SOP support.

We also contribute more generally to the literature on readability of corporate narrative disclosures. To our knowledge, our paper is the first to demonstrate heterogeneity in the effects of readability when accounting for investors’ prior concerns, e.g., apparent pay excesses in our setting. This suggests an important priming role of trust in conditioning the effect of readability on perceptions formed when reviewing disclosures. Further exploration of conditionality in

shareholders' responses to other disclosures (e.g., earnings announcements, periodic reports) according to, e.g., indicators of agency concerns may be a fruitful avenue for future research.

The remainder of the paper is organised as follows: Section 2 introduces the institutional setting for the SOP and related compensation disclosure within the US context, and presents the development of our hypotheses. The sample data and construction of our main variables is described in Section 3. The empirical results are presented in Section 4, while Section 5 concludes the paper.

2. Background, Prior Literature and Hypotheses

2.1 Institutional Setting

Although requirements for enhanced compensation disclosure within the annual proxy statement preceded Say-on-Pay (SOP) by several years, and advisory SOP votes were held by certain US corporations prior to 2011, the Dodd-Frank Act (2010) mandated that all large public US corporations hold a non-binding shareholder vote from the 2011 proxy season onwards, and at least once every three years. Section 14A of the Dodd-Frank Act also requires companies to allow shareholders to vote on the frequency of the SOP, and Ernst and Young (2018) report that 90% of Russell 3000 companies now provide for an annual vote.

To inform the SOP, the SEC requires a CD&A section to be included in the DEF 14A proxy statement, and that it be written in Plain English.³ The CD&A is published after the end of the fiscal year to which the SOP relates, typically 30 to 45 days before the AGM during which the vote will be held. The SEC (2006, p. 191-192) suggest that “*clearer, more concise presentation of executive and director compensation matters... can facilitate more informed investing and*

³ The SEC (2006, p. 192) advocate, for example, the use of “*short sentences*”, and “*definite, concrete, everyday words*”, consistently with our use of the Gunning-Fog Index to measure readability, which considers readability to be inversely related with average sentence length and the proportion of complex words.

voting decisions in the face of complex information". The SOP has thus been viewed as a catalyst for shareholder engagement and dialogue, with failures in "*shareholder outreach and disclosure*" noted by Semler Brossy (2016, p. 5) as a common contributory factor in cases of high SOP dissent.

Although the SEC's Plain English initiative was primarily intended to benefit the "*least sophisticated investors*" (SEC, 1998), the previous SEC Chairperson, Christopher Cox reflected that "*the average Compensation Disclosure and Analysis section isn't anywhere close to Plain English*", suggesting that "*most of it is as tough to read as a Ph.D. dissertation*" (Cox, 2007). Larcker et al. (2015) also found only 38% of institutional investors surveyed consider the typical compensation disclosure in the proxy statement to be clear and easy to understand.

Whilst the SOP vote is non-binding, levels of dissent of above 25%-30% have been found to place pressure on boards to address shareholder concerns.⁴ Ertimur et al. (2013) find that 32% of firms receiving just *under* 30% SOP dissent instil compensation changes, but that this increases discontinuously to 72% for firms receiving just *over* 30% dissent. Relatively few firms (approx. 10%) receive SOP dissent of 30% or more, however, and in the majority of cases dissent is below 10% (Brunarski et al., 2015). The academic literature typically considers 'high' dissent as anything over 30% (e.g., Brunarski et al., 2015) or in some cases 20% (Correa and Lel, 2016).

Hooghiemstra et al. (2017), amongst others, provide useful insights regarding the UK context but the setting of the US SOP differs in several important respects. Firstly, while UK remuneration disclosures are written by the board of directors, responsibility for the CD&A lies with senior executives. As it is also senior executives' remuneration that the US SOP

⁴ The two largest proxy advisory services, Glass Lewis and ISS, have policies to place increased levels of scrutiny on firms receiving greater than 25% and 30% dissent, respectively.

relates to, potential agency problems, specifically incentives to produce more persuasive remuneration disclosures, are arguably heightened in the US context. Secondly, broker voting of shares held in street name has been disallowed in the US since the inception of the SOP but is permitted in the UK. This is important as brokers tend to cast votes disproportionately in favour of management (Bethel & Gillan, 2002; Cai et al., 2017). Thirdly, CD&A of US firms are typically longer and more complex than DRR. For example, the average (median) length of US CD&A in our sample is 6,814 (6,830) words. Data compiled by the Corporate Financial Information Environment project (El-Haj et al., 2019a) indicates that the average (median) length of UK DRR over the same period was 3,998 (2,513) words.⁵

2.2 Writing the CD&A – The Preparers’ Perspective

Whilst our main focus is on the users’ perspectives, i.e., whether readability affects investors’ judgements, it is important to also consider the preparer’s perspective, since CD&A may be prepared strategically, to increase shareholder approval.

There is some evidence that managers exploit investors’ limited ability to process information by strategically reducing readability in order to obfuscate bad news. For instance, Li (2008) finds evidence that companies use strategic obfuscation in their annual reports to hide poor results. Warren Buffett comments, in the preface to the SEC’s (1998, p. 1) Plain English Handbook, *“In some cases, ... I suspect that a less-than-scrupulous issuer doesn’t want us to understand a subject it feels legally obligated to touch upon”*. Laksmana et al. (2012) reports evidence of obfuscation of CEO pay when CD&As were first introduced and Robinson et al.

⁵ While the firms in our sample are likely to be, on the whole, larger than the average UK firm, this is unlikely to explain the pronounced disparity in typical remuneration disclosure length. All else equal, we observe each additional \$1bn of sales to be associated with CD&A that are approximately 8 words longer on average. Mean (median) sales in our sample is \$12bn (\$3bn) implying an approximate 96 (24) word increment in typical word-count, over the very smallest firms. Moreover, the mean (median) word length of UK DRR situates just above (below) the 10th percentile of our US CD&A sample.

(2011) document that non-compliance in the first wave of CD&A, as identified by SEC critiques, was positively associated with excess CEO compensation.⁶ Hooghiemstra et al. (2017) also provide evidence of obfuscation in UK Directors' Remuneration Reports (DRR).

Bloomfield (2008) intuitively suggests that readability is also affected by 'ontology' (i.e., the nature of subject matter). Firms may therefore naturally default to highly technical language if the effort required to write in Plain English is not considered worthwhile. Bushee and Miller (2012) report that smaller, less visible firms, in particular, prefer to develop direct channels of communications with target investor groups rather than improve disclosure. Improved disclosure may even induce a more volatile stock price by attracting transient investors (Bushee and Noe, 2000), an outcome which some companies may be keen to avoid.

2.3 Effects of CD&A Readability – The Users' Perspective

A substantial part of an investor's information set regarding the executive compensation package is conveyed through narrative disclosure within the CD&A. Since investors' attention is a scarce resource (Hirshleifer and Teoh, 2003), the costs of collecting information from the typical CD&A are likely to be high. Considerable time and effort is required to process and interpret complex communication (Lehavy et al., 2011). Bloomfield (2002) postulates that investors' judgments are less informed when information collection costs are high since collection costs can outweigh the perceived benefits.

Consistent with this hypothesis, Lehavy et al. (2011) find less readable 10-K reports are associated with greater dispersion and lower accuracy in analyst forecasts, suggesting even

⁶ CD&A readability may be affected by the risk of litigation, if increased involvement of legal council in drafting compensation disclosure leads to a greater prevalence of boilerplate language (Serafin, 1998). The second SOP proxy season (2012), in particular, saw significant numbers of lawsuits related to proxy compensation disclosures (Ross, 2013). Evidence suggests that stronger legal compliance of disclosures can mitigate litigation risk (Hanley and Hoberg, 2012).

sophisticated audiences are less informed when readability is poor.⁷ Lawrence (2013) argues that difficult-to-read disclosures may place retail investors, in particular, at an informational disadvantage and suggests individual investors invest less in firms with lower quality disclosures.

Rather than less-readable remuneration disclosures not affecting voting outcomes, Hooghiemstra et al. (2017) find that obfuscation, due to lower readability, results in lower SOP dissent. Poor readability may leave shareholders confused, and less able to form unambiguous conclusions on the appropriateness of CEO pay. In the absence of clear reasoning to believe that pay arrangements are inappropriate, shareholders left confused may simply revert to a ‘default’ position of endorsing the pay arrangements. In other words, shareholders may be less likely to challenge pay arrangements if they do not fully understand the content of the CD&A (Hooghiemstra et al., 2017). We therefore state our first hypothesis, in its alternative form and according to the obfuscation hypothesis, as follows:

H1 Lower CD&A readability decreases the probability of dissent in the absence of clear excesses in CEO pay.

A seemingly contradictory perspective to the obfuscation hypothesis is that shareholders may become more sceptical, rather than dismissive, of pay arrangements as a result of poor readability. Rennekamp (2012), amongst others, suggest investors find disclosures that are more difficult to process to be less credible and suggest processing fluency (high readability) can, subconsciously, improve confidence in the message conveyed. Hwang and Kim (2017, p. 374) conjecture that “*lower readability undermines investors’ belief that a source can be trusted*”. Shah and Oppenheimer (2007) also suggest processing fluency increases individuals’

⁷ Moreover, while the market typically under-reacts to 10-K filings, according to You and Zhang (2009), this is even more so for longer reports. Similarly, Lee (2012) finds that under-reaction to earnings announcements (the post-earnings announcement drift) is exacerbated when the corresponding 10-Q report is abnormally difficult to process.

reliance on information (such as management's justification of executive pay arrangements in the CD&A) when making decisions. Rennekamp (2012) further argues that, even when the same information is being conveyed, investors evaluate the credibility of management to be lower when processing fluency (readability) is poor. The processing fluency hypothesis would therefore predict that shareholders are more likely to consider excessive CEO pay to be unjustified when the CD&A is less readable.

As previously mentioned, SOP outcomes are particularly sensitive to excesses in CEO compensation levels. Low pay-performance sensitivity, i.e. proposed pay levels that are unwarranted given corporate and/or stock performance, can be a significant predictor of SOP dissent (Collins et al., 2019). It is implausible that shareholders would vote 'against' when CEO pay is substantially below the benchmark (i.e., when the CEO is underpaid) purely based on a difficult to read CD&A. On the other hand, we expect that scepticism of pay arrangements is likely to be heightened when CEO pay is highly excessive. For example, Stathopoulos and Voulgaris (2016) demonstrate that less-engaged shareholders refrain from casting a negative vote unless CEO pay is clearly excessive. We therefore conjecture that while shareholders' default position may be to *endorse* pay arrangements when CEO pay is at normal levels, the default is more likely to be to *challenge* pay arrangements when CEO pay prima-facie appears highly excessive (e.g., is well above benchmarks).

Shareholders will perceive a more pressing need for justification when CEO pay appears highly excessive, and therefore a clear understanding of the rationale for pay arrangements (facilitated through greater Plain English) would be pivotal in quelling their concerns. Thus, while obfuscation may be successful when excess pay is at moderate levels, we expect that shareholders will become more (not less) sceptical of pay arrangements as a result of low CD&A readability, when CEO pay is clearly excessive. Specifically, we conjecture that CD&A readability moderates the effect of *Excess comp* on *Dissent*, such that less readable CD&A

undermine shareholders' perceptions of management credibility, and tolerance of excessive pay, when compensation levels are clearly high. Low CD&A readability is thereby predicted to strengthen the impact of *Excess comp* on the probability of *Dissent*.⁸ We therefore state our second hypothesis, based on the processing fluency hypothesis, as follows:

H2 Lower CD&A readability increases (positively moderates) the conditional probability of dissent given the extent of excess CEO pay.

2.4 The Presence of Institutional Shareholders

Engagement in monitoring activities varies depending on the incremental benefits and costs faced by each investor (Stathopoulos and Voulgaris, 2016). The collection of information from corporate documents is costly in terms of time and effort (Bloomfield, 2002), thus anticipation of greater derived benefits from a thorough review of the CD&A is necessary to render the effort worthwhile. Consistent with this, prior literature demonstrates that the intensity of monitoring is greater when ownership is characterised by large institutional shareholders who are able to benefit from economies of scale and hence face lower proportional information collection costs (Chen et al., 2007). Sophisticated investors (e.g., institutional investors) may also have a comparative advantage, since, according to Bushman et al. (1996, p. 52), unsophisticated investors are more likely to suffer from “*information overload*”.

There is evidence that large institutional investors actively engage in effective monitoring of executive pay. For example, Hartzell and Starks (2003) find that pay-for-performance sensitivity is positively related to institutional ownership concentration, while Khan et al. (2005) observe that a greater proportion of equity held by the largest institutional investor is

⁸ Heterogeneous effects of CD&A readability on *Dissent* may explain why Lo et al. (2014) report evidence of a negative relationship between CD&A readability and SOP dissent, Hooghiemstra et al. (2017) provide evidence of a positive relationship, while Balsam et al. (2016) find no significant effect. These prior studies do not consider any moderating effect of readability on the link between *Excess comp* and *Dissent*.

associated with more moderate CEO pay.⁹ On the other hand, Stathopoulos and Voulgaris (2016) observe that short-term oriented institutional investors, who face lower monitoring incentives, are more likely to avoid expressing opinion by abstaining from the SOP altogether. Hooghiemstra et al. (2017) also find that lower DRR readability is associated with greater levels of SOP dissent for firms with high levels of institutional ownership.

For small investors, the anticipated costs of processing the CD&A may outweigh the perceived benefits altogether, hence they may choose to ignore the CD&A, in which case, *variations* in CD&A readability would have little impact on SOP voting patterns. Thus, the impact of readability on the SOP may be stronger for companies with greater levels of institutional ownership, i.e., when monitoring intensity, and thus incentives for engagement with the CD&A, are high.

Conversely, according to Tan et al. (2014), variations in CD&A readability may have a greater impact on small shareholders, providing they *engage* with the disclosure, as they are more likely to be influenced by heuristic processing such as the processing fluency heuristic. Large institutional investors, however, are more likely to engage in systematic information processing, rather than heuristic processing, due to greater monitoring incentives as well as a greater level of sophistication. This suggests that institutional investors' judgements will be less influenced by variations in CD&A readability. While we discuss above that the effects of CD&A readability on small shareholder voting (e.g. the processing fluency heuristic) are conditional on small shareholders engaging fully with the SOP, potentially unfair executive remuneration is a topic which attracts considerable capital market and public outrage, and investment professionals often have relatively higher outrage thresholds (Arnold & Grasser,

⁹ Evidence from the securities lending market also suggests that large institutional investors value their ability to express voice through voting, restricting lendable supply and even recalling loaned shares prior to proxy record dates (Aggarwal et al., 2015). Such reductions in equity lending are found to be greater for firms with high concentrations of institutional investors.

2018). Thus, there are strong reasons to expect that even small shareholders will engage with the process, and we therefore state our third hypothesis as follows:

H3 The effect of CD&A readability on the conditional probability of dissent, given the level of excess compensation, is greater for firms with low institutional ownership.

3. Data and Key Variables

3.1 Sample Selection

We construct our sample from the population of S&P 1500 firms holding a shareholder SOP vote for fiscal years 2010 to 2014.¹⁰ We collect DEF 14A proxy statements from the SEC EDGAR database and isolate the CD&A following the parsing procedure outlined in Appendix B. We obtain SOP voting results from Bloomberg,¹¹ as well as CEO-level data. Firm-level financial data are from Worldscope, and returns data are from Datastream.¹² Our final sample consists of 2,686 firm-year observations with all necessary data available. Of these, there are 219 instances of firms receiving greater than 30% dissent, or 8.15% of the full sample. This compares very closely to frequencies of >30% dissent for S&P 1500 firms observed by Ertimur

¹⁰ We retain financial firms in the sample alongside non-financial firms, as controversy regarding executive compensation is equally manifest in this sector. We observe no significant differences in the results between financial and non-financial firms.

¹¹ We check the integrity of the voting data by also hand-collecting results from 1,000 Form 8-K filings. We observe an agreement rate of 98.8% with voting data obtained from Bloomberg.

¹² We use accounting data from the Worldscope database as Ulbricht and Weiner (2005) find no statistical or methodological shortcomings in comparison with Compustat for US firms. They also report that since 1998 a greater coverage of firms is provided by Worldscope, though as we focus on S&P 1500 firms any difference in coverage is likely to be minimal. Overall, they conclude that use of both databases should lead to comparable results. While some papers suggest greater caution should be used when using returns data from Datastream as opposed to CRSP (Ince & Porter, 2006), Schmidt et al. (2019) find that aggregate market returns and risk factors calculated using Datastream data compare very closely to those derived from CRSP. Ince and Porter (2006) identify potential security-level issues in Datastream, however, including incorrect identification of the exchange of listing. As they state that most of the problems are concentrated in smaller stocks and/or low-price stocks, we do not expect our analyses of S&P 1500 firms to be significantly affected by our database choices. As we employ yearly returns, potential short-term biases are likely to be smoothed-out. Similar to prior papers, we also winsorise at the 1st and 99th percentiles to mitigate any remaining issues. A number of prior studies choose to employ Datastream for U.S. returns data due to the breadth and depth of its coverage (Hou et al. 2011).

et al. (2013) and Brunarski et al. (2015) of 8.2% and 8.88%, respectively. For further tests, we supplement our dataset with institutional ownership data from Thomson One Banker. We obtain ownership data for 2,141 firm-years (173 of which have greater than 30% dissent). We employ the full sample of 2,686 observation in tests that do not require institutional shareholdings data.

3.2 Estimating Excess CEO Compensation

We estimate a baseline measure of excess CEO compensation (*Excess comp*) following Core et al. (2008), since SOP outcomes are shown to be particularly sensitive to the excessiveness of CEO compensation (Brunarski et al., 2015). We focus on excesses in terms of total CEO compensation (*Total comp*), being the sum of salary, bonus, long-term incentive plan pay-outs, the value of restricted stock grants, the value of options granted during the year, and any other annual pay; as reported in the summary compensation table of the DEF 14A. *Excess comp* is estimated as the residuals from cross-sectional regressions of $\ln(\text{Total comp})$ on a range of firm-specific variables. That is, *Excess comp* is equal to $\ln(\text{Total comp})$ minus its fitted value (the ‘expected’ level of pay). Specifically, we run rolling OLS regressions predicting *Excess comp* in year t using observations from years $t-2$ to t , taking the following form:

$$\begin{aligned} \ln(\text{Total comp}_{i,t}) = & \alpha + \beta_1 \ln(\text{tenure}_{i,t}) + \beta_2 \ln(\text{sales}_{i,t}) + \beta_3 \text{S\&P500}_{i,t} \\ & + \beta_4 \text{Bk/Mkt}_{i,t} + \beta_5 \text{ROA}_{i,t} + \beta_6 \text{ROA}_{i,t-1} + \beta_7 \text{RET}_{i,t} + \beta_8 \text{RET}_{i,t-1} \\ & + \gamma \cdot \text{year} + \delta \cdot \text{industry} + \varepsilon_{i,t} \end{aligned} \quad (1)$$

Where *tenure* is CEO tenure at the year-end in years; *sales* is firm sales at the end of year t ; *S&P500* is an indicator equal to 1 if the firm is a constituent of the S&P 500, and 0 otherwise; *Bk/Mkt* is the book-to-market ratio calculated as (book value of assets)/(book value of liabilities + market value of equity); *ROA* is return on assets (operating income scaled by total assets);

and *RET* is the stock return over the 12 months to the year-end.¹³ We also include year and industry fixed effects, based on the Fama-French 12 industry classification. Year t is the fiscal year for which *Total comp* is awarded (the year the SOP vote relates to, as opposed to the year in which the vote takes place).

Excess comp, in year t , is taken as the residual from Eq. (1), which orthogonalizes CEO pay to cross-sectional predictors during years $t-2$ to t . Cumulatively, our regressions therefore include data for S&P 1500 firm-years from 2008 to 2014. While Core et al. (2008) model expected compensation using annual cross-sectional regressions, Brunarski et al. (2015) run a single pooled regression employing four years' data. Though we estimate *Excess comp* using three-year rolling regressions, for brevity we present estimates for the full 2008-2014 period in Table 1. Similar to Core et al. (2008), we find that $\ln(\textit{Total comp})$ is significantly higher for CEOs of larger firms, S&P500 constituents, and firms realising higher stock returns over the last two years, but is unrelated to $\ln(\textit{tenure})$. Our results somewhat differ from Core et al. (2008), however, as CEO pay does not appear to be negatively related to *Bk/Mkt* or *ROA* for either of the prior two years. We find firm size, measured by the log of *sales*, to be the most significant predictor of CEO compensation levels.

[Insert Table 1 about here]

3.3 CD&A Readability

We focus on the readability of the CD&A section of the proxy statement, specifically, since:

1. DEF 14A text outside of this section mostly relates to matters other than executive remuneration;

¹³ We also include $\ln(\textit{Total assets})$ instead of $\ln(\textit{Sales})$ in Eq. (1) and obtain consistent results.

2. Remuneration disclosures outside of the CD&A mainly consist of tables and technical notes which are not amenable to text analysis; and
3. The purpose of the CD&A is to “*plainly tell the company’s compensation story*” (Cox, 2007). As such, the SEC specifically require the CD&A to be written in Plain English.

Our primary measure of CD&A readability is the Bog Index, as suggested by Bonsall et al. (2017). The Bog Index is a readability measure provided by the computational linguistics software, *StyleWriter – The Plain English Editor*. The main advantage of the Bog Index, as a measure of CD&A readability, over alternative measures, is that it is based on a range of Plain English attributes recommended in the SEC’s (1998) Plain English Handbook. The Bog Index is calculated as the sum of three broad components: (1) Sentence Bog; (2) Pep; and (3) Word Bog. Sentence Bog reflects the diminution of readability from longer sentences, in a non-linear fashion (it is based on the squared average sentence length). Pep reflects attributes that facilitate readability, by making the writing more interesting (e.g., use of names and interesting words). Finally, Word Bog reflects a range of word-level attributes which affect readability either due to (1) Plain English style problems (e.g., passive verbs, legal terms, clichés, etc.), and/or (2) word difficulty (based on a proprietary dictionary).

The Bog Index is an inverse measure of readability, with high Bog Index values reflecting low readability, or a low level of Plain English. The *StyleWriter* software defines the readability of texts with Bog Index values above 70 as ‘poor’, above 100 as ‘bad’, and above 130 as ‘dreadful’. We estimate the Bog of CD&As in our sample to have a mean of 96.5 (median = 96), with a range between 60 and 177, thus indicating the average readability of CD&A disclosures in our sample to be ‘poor’, and a significant portion (31.2%) are defined as either ‘bad’ or ‘dreadful’.

We also estimate the readability of the CD&A using the Gunning-Fog Index (*Fog*).¹⁴ The Fog Index measures textual complexity as a function of (a) the number of words per sentence, and (b) the percentage of words identified as ‘complex’. Complex words are defined as those consisting of three or more syllables. Eq (2), below, describes the formula for calculating the Fog Index. Since the Fog Index measures textual complexity, higher values of the Fog Index indicate less readable disclosures. A standard interpretation of Fog Index values is that they denote the number of years of formal education necessary in order for a reader to clearly understand a text on a first reading.

$$\text{Fog Index} = 0.4 \left[\left(\frac{\text{No. words}}{\text{No. sentences}} \right) + 100 \left(\frac{\text{No. complex words}}{\text{No. words}} \right) \right] \quad (2)$$

The Fog Index originates from the computational linguistics literature but it has been widely used in Accounting and Finance studies examining readability. Recent examples include Li (2008), Lehavy et al. (2011), Dougal et al. (2012), Li and Zhang (2015), and Hsieh et al. (2016). We therefore provide results based on the Fog Index to facilitate comparability with related studies.

The Fog Index, and in particular, the determination of ‘complex’ words, is not without criticism. The SEC’s (1998, p. 30) guidance on Plain English advocates the use of “*short, common words*” in place of jargon, even those that are likely to be familiar to industry experts.¹⁵ Rennekamp (2012) suggests that simplification of text in this way facilitates information processing, while not necessarily affecting ultimate comprehension of information communicated. Nevertheless, Loughran and McDonald (2014, p.1645) criticise the use of the

¹⁴ We also obtain consistent results when measuring readability as the average length of sentences, and when using the Flesch-Kincaid Grade Level formula.

¹⁵ The SEC (1998, p. 31) suggest the following: “*Surround complex ideas with short, common words. For example, use **end** instead of **terminate**, **explain** rather than **elucidate**, and **use** instead of **utilize**. When a shorter, simpler synonym exists, use it*” (emphasis in the original). In each of these examples, readability is facilitated by using words of one or two syllables rather than alternatives of three syllables or more.

Fog Index to measure the readability of financial disclosures, arguing that frequently used words in a financial context, such as ‘financial’ and ‘management’, are treated as difficult to understand when they are not. By contrast, the Bog Index incorporates a more reliable measure of word difficulty, based on a proprietary dictionary rather than the number of syllables.

Li (2008) suggests that a Fog Index value of 18 or greater indicates that a text is unreadable. We estimate the *Fog* of CD&As in our sample to have a mean of 23.6 (median = 23.4), with a range of between 18.3 and 43.5 suggesting every CD&A in our sample would be considered unreadable. From the standard interpretation of *Fog*, our data suggests that *at least* 18.3 years of formal education are required in order to comprehend a CD&A on first reading. We also observe considerable persistence in readability over time, with a serial-correlation of 0.74 (0.79) between the *Bog (Fog)* of CD&A in successive years.¹⁶

3.4 Institutional Ownership

We collect details of institutional investor ownership, at the end of the most recent quarter prior to the SOP vote, from Thomson One Banker, and measure total institutional ownership as the proportion of outstanding shares held by all 13-F institutional investors (*%Inst*).¹⁷ Since monitoring is likely to involve fixed costs in terms of information processing, greater monitoring activity is expected in the presence of large institutional shareholders, who may benefit from economies of scale (Chen et al., 2007).

3.5 Descriptive Statistics

A full set of descriptive statistics is provided in Table 2. We observe the mean (median) proportion of dissenting votes to be relatively low, at just over 9% (4%), consistent with

¹⁶ Due to the strong persistence in the readability of firms’ CD&A, we cluster standard errors at the firm-level in our regression analyses.

¹⁷ We also obtain consistent results when employing the proportion of shares held by institutional blockholders (institutions owning at least 5% of outstanding shares).

samples examined by prior studies (Brunarski et al., 2015). Since this variable is highly skewed, we consider its use as a continuous dependent variable, within a conventional OLS regression framework, to be potentially problematic. Moreover, Brunarski et al. (2015) advocate focusing on the conditional *probability* of receiving greater than 30% Dissent (receiving lower than 70% support), since firms exceeding this threshold, in particular, are subject to significant increased scrutiny. Consistent with this view, Ertimur et al. (2013) demonstrate that firms' responsiveness to SOP dissent increases discontinuously as the proportion of dissenting votes exceeds 30%.¹⁸

[Insert Table 2 about here]

A correlation matrix for our main variables is presented in Table 3. We observe a slight positive correlation between *Bog* and *Excess comp* ($\rho=0.09$), and *Fog* and *Excess comp* ($\rho=0.06$), consistent with deliberate obfuscation. The correlation between *Bog* and *Dissent*, and *Fog* and *Dissent*, however is slightly positive ($\rho=0.04$ and 0.06 , respectively), rather than negative, which is inconsistent with the obfuscation hypothesis that low readability reduces *Dissent*. A more substantive correlation is found between *Excess comp* and *Dissent* ($\rho=0.35$), consistent with prior findings that *Dissent* is particularly sensitive to excesses in CEO pay (Brunarski et al, 2015).

[Insert Table 3 about here]

We visually examine the three-way relationship between *Dissent*, *Excess comp*, and *Bog* (*Fog*) further in Figure 1, Panel A (Panel B). In Panel A of Figure 1 we plot the frequency of firms obtaining dissenting SOP outcomes (greater than 30% of votes cast against) across 25 groups based on quintiles of *Bog*, and quintiles of *Excess comp* (5x5 groups). In Panel B, we substitute

¹⁸ For robustness, we re-estimate our models using OLS regressions, and employ continuous measures for SOP Dissent (%) and our readability measures. The results are presented in Section 4.3.5.

Fog instead of *Bog* as the readability measure. Intuitively, the frequency of *Dissent* is greater when *Excess comp* is in the highest quintile, however the frequency of *Dissent* for firms with *Quint5 Excess comp* increases near-monotonically, and substantially, across *Bog/Fog* quintiles. This lends indicative support to hypothesis 2, that lower CD&A readability strengthens the link between excess CEO pay and SOP dissent. Of firms with the most over-paid CEOs (*Quint5 Excess comp*) and also the least readable CD&A (either *Quint5 Bog* in Panel A or *Quint5 Fog* in Panel B), we observe 42 instances of *Dissent*. This compares with only 13 and 8 instances, respectively for firms with equally over-paid CEOs (*Quint5 Excess comp*) but produce among the most readable CD&A, indicated by *Quint1 Bog* and *Quint1 Fog*.

[Insert Figure 1 about here]

3.6 Empirical Model

To examine the potential moderating effect of CD&A readability on the link between *Dissent* and *Excess comp*, we estimate the following logit regression model:

$$\begin{aligned}
Dissent_{i,t} = & \alpha + \beta_1 Excess\ comp_{i,t} + \beta_2 High\ Bog_{i,t} + \beta_3 Excess\ comp_{i,t} * High\ Bog_{i,t} \\
& + \beta_4 \ln(tenure_{i,t}) + \beta_5 IVOL_{i,t} + \beta_6 Neg_RET_{i,t} + \beta_7 RET_{i,t} + \beta_8 Neg_RET_{i,t} * RET_{i,t} \\
& + \beta_9 Loss_{i,t} + \beta_{10} ROA_{i,t} + \beta_{11} Loss_{i,t} * ROA_{i,t} + \beta_{12} EDec_{i,t} + \beta_{13} dROA_{i,t} \\
& + \beta_{14} EDec_{i,t} * ROA_{i,t} + \beta_{15} \ln(sales_{i,t}) + \gamma \cdot year + \delta \cdot industry + \varepsilon_{i,t}
\end{aligned} \tag{3}$$

Where *Dissent* is an indicator variable equal to 1 if greater than 30% of votes are cast ‘against’ proposed executive pay in the SOP, and 0 otherwise; *Excess comp*, a measure of the excessiveness of CEO pay, as described in Section 3.2, is measured as the residual from Eq. (1); *High Bog* is an indicator variable equal to 1 if the Bog Index value of the CD&A is above the median for all observations in our sample, and 0 otherwise; and *IVOL* is the idiosyncratic volatility of stock returns over the 12 months to the fiscal year-end, calculated as the Root Mean Squared Error (RMSE) from the Fama-French 3 factor model. *RET*, *ROA*, and *ln(sales)*

are as previously defined. We also include *Neg_RET*, equal to 1 if *RET* is negative, and 0 otherwise; *Loss*, equal to 1 if *ROA* is negative, and 0 otherwise; *dROA*, the percentage change in *ROA* between years *t-1* and *t*; and *EDec* (earnings decline), equal to 1 if *dROA* is negative, and 0 otherwise. We include interactions between *Neg_RET* and *RET*, *Loss* and *ROA*, and *EDec* and *dROA* to control for possible non-linear effects of performance on *Dissent*.¹⁹ In some tests we substitute *High Bog* with *High Fog*, an indicator equal to 1 for observations with above-median values of Fog Index, and 0 otherwise.

Our main coefficients of interest are β_2 on *High Bog* (*High Fog*), which serves as the basis for testing hypothesis 1, and β_3 on the interaction between *Excess comp* and *High Bog* (*High Fog*), which serves as the basis for testing hypothesis 2.²⁰ According to the obfuscation hypothesis, lower readability reduces the probability of *Dissent* (i.e., $\beta_2 < 0$). On the other hand, the processing fluency hypothesis predicts lower readability will strengthen the link between *Excess comp* and the probability of *Dissent* (i.e., $\beta_3 > 0$).

4. Empirical Results

4.1 Predicting Say-on-Pay Dissent

We first examine the conditional probability of receiving greater than 30% voting dissent, *Dissent*, given *Excess comp*, *High Bog* (or *High Fog*), and their interaction, within a logistic regression framework using the full sample. The results are presented in Table 4. In models (1) and (2) we estimate Eq. (3) using the full sample, while in models (3)-(4) and (5)-(7) we run estimations using underpay and overpay sub-samples. The overpay sub-sample comprises observations with positive values for *Excess comp*, whereas the underpay sub-sample

¹⁹ Shareholder dissatisfaction with management may increase discontinuously if they fail to meet basic performance thresholds, e.g., if earnings or stock returns over the year are below zero.

²⁰ We are mindful of the empirical challenges in interpreting coefficients on interactions from logistic regressions. We therefore employ graphical analyses to supplement our modelling, which are discussed in more detail in the next section.

comprises observations with negative values. In the interest of brevity, we present results using *High Fog* for the overpay sample only, i.e., model (7). In models (1), (3), and (5) we restrict β_2 and β_3 to zero. In each of the alternative specifications presented in Table 4, with the exception of model (4) which examines only the underpay sample, the estimated coefficient on *Excess comp* is positive and highly significant. Thus, consistent with Brunarski et al. (2015), we observe that excessive CEO pay significantly increases the likelihood of *Dissent*.

[Insert Table 4 about here]

The estimated coefficients on the control variables are generally as expected, with a higher probability of *Dissent* for firms with lower stock returns (*RET*), lower profitability (*ROA*), and earnings declines (*EDec*), and for firms with longer-serving CEOs (log of *Tenure*). Firms with longer-serving CEOs having a higher probability of *Dissent* is consistent with the notion that CEO entrenchment leads to greater agency problems and Correa and Lel's (2016) finding that SOP helps to address these problems. Holding all else equal, we also find that larger firms are more likely to receive *Dissent*. It is possible that this reflects the concentration of public scrutiny on the highest paid CEOs, since $\ln(\text{Total comp})$ is strongly correlated with firm size as measured by the log of *Sales* ($\rho = 0.73$). *Excess comp* is, by construction, orthogonal to the pronounced size effect that explains much of the variation in *Total comp*; i.e., it captures *abnormal* pay, as opposed to *high* pay, per se.

In models (2), (4), and (6) we include *High Bog*, as well as the interaction with *Excess comp*. We find the main effect of *High Bog* on *Dissent* to be negative but insignificant in each case. Though insignificant, a negative coefficient on *High Bog* is directionally consistent with obfuscation at low levels of *Excess comp*. We observe a positive and significant coefficient however on the interaction between *Excess comp* and *High Bog* in the full sample (model 2) and the overpay sample (model 6). A positive estimated coefficient of 0.738 on the *Excess*

*comp*High Bog* interaction in the overpay sample (model 6, t-stat = 1.98) suggests that the sensitivity of *Dissent* to *Excess comp* is more pronounced for firms with less readable CD&A (*High Bog* = 1). The coefficient on *Excess comp* also reduces from 2.302 in model (5) to 1.891 in model (6), consistent with the notion that the sensitivity of *Dissent* to *Excess comp* is lower when CD&A are more readable (*High Bog* = 0).

We present results using *High Fog* in place of *High Bog* in model (7), using the overpay sample. We observe a significantly negative coefficient for *High Fog* and a significantly positive coefficient for the *Excess comp*High Fog* interaction variable. At low levels of *Excess comp*, the negative coefficient on *High Fog* dominates suggesting that low CD&A readability results in lower dissent due to obfuscation. At higher values of *Excess comp*, however, the increased likelihood of *Dissent* given *Excess comp* due to low readability (indicated by a significantly positive β_3), outweighs the negative level effect of *High Fog*. This suggests that a critical value of *Excess comp* exists, above which the obfuscation effect of low readability gives way to increased scepticism (processing fluency effects).

We also present at the foot of Table 4 the average marginal effect of *Excess comp*, as well as the average marginal effect of *Excess comp* when *High Bog* or *High Fog* is 0 (1) separately. The statistics support the above interpretation of our coefficient estimates. In particular, in relation to model (6) we observe that a unit increase in *Excess comp* is associated with a 16% increase in the probability of *Dissent* when *High Bog* is 0 but a 22.3% increase when *High Bog* is 1. In relation to model (7) we observe that a change in *High Fog* from 0 to 1 is associated with a doubling of the average marginal effect of *Excess comp* from 12.4% to 25.2%.

Interpretation of the coefficients on interaction terms in logit regressions, however, is problematic, since the magnitude and direction of the interaction effect can differ to that of the estimated coefficient (Ai & Norton, 2003). Moreover, according to Greene (2010, p. 295)

statistical testing about partial effects of interaction terms in non-linear models “*produces generally uninformative and sometimes contradictory and misleading results*”, and graphical analyses are recommended as an addendum. We therefore follow Greene (2010) and De Jong et al. (2012) in producing graphical displays where values of the interacted variables are plotted against the predicted probability given the estimated model.

In Panel A (Panel B) of Figure 2 we display the predicted probability of Dissent in relation to values of the interacted variables *Excess comp* and *High Bog (High Fog)*, based on model 6 (model 7) of Table 4. We plot two lines, one for the case when *High Bog (High Fog)* = 0, and another for the case when *High Bog (High Fog)* = 1. All other variables in the model are held at their mean value. In each case, the lines denote the predicted probability of *Dissent* over a range of values of *Excess comp*. Divergence between the two lines thereby depicts the impact of the interaction effect on the probability of *Dissent*. For ease of interpretation, we express *Excess comp* in terms of standard deviations away from the mean, which, by construction, is approximately zero.

[Insert Figure 2 about here]

In Panel A (Panel B) of Figure 2 we observe that the interaction between *Excess comp* and *High Bog (High Fog)* serves to widen the distance between the plotted predicted probabilities. This indicates that the positive effect of *Excess comp* on the probability of *Dissent* is greater when *High Bog (High Fog)* = 1 than when *High Bog (High Fog)* = 0, consistent with hypothesis 2. According to the Bog readability measure in Panel A, the probability of *Dissent* when *Excess comp* is two standard deviations greater than the mean is 27.5% when *High Bog* = 1, but much lower at 17.8% when *High Bog* = 0. Strikingly, according to the *Fog* readability measure in Panel B, the probability of *Dissent* when *Excess comp* is two standard deviations greater than

the mean is over twice as high when the readability of the CD&A is low (33.9% when *High Fog* = 1 versus 13.8% when *High Fog* = 0).

4.2 The Role of Institutional Monitoring

Hypothesis 3 predicts that the effect of CD&A readability on the conditional probability of *Dissent* given *Excess comp* is greater for firms with low levels of institutional ownership. In Table 5, we estimate Eq. (3) using sub-samples of firms with high and low levels of institutional shareholdings. High (low) institutional ownership (IO) comprises firms with above (below) median values of *%Inst* (percentage of shares outstanding held by institutional investors). Results for the full sample are presented in models (1) and (2), while results for the overpay sample are presented in models (3)-(6). We present results using *Bog* in models (1)-(4), and results using *Fog* in models (5)-(6).

[Insert Table 5 about here]

In models (1) and (2), we observe that the coefficients on *High Bog* and *Excess comp*High Bog* become insignificant when splitting the full sample by high/low institutional ownership. In models (3) and (4), however, where we examine only the overpay sample, we observe a significantly positive coefficient on *Excess comp*High Bog* in the low IO sub-sample (model 4), but not the high IO sub-sample (model 3). We conclude, therefore, that our findings lend support for hypothesis 3, that the effect of CD&A readability on the conditional probability of *Dissent*, given *Excess comp*, is greater for firms with more unsophisticated shareholders. This is consistent with less sophisticated investors relying more on heuristic processing, rather than systematic processing. Consistent with this conjecture, we also find some evidence of obfuscation in the low IO sub-sample (model 4) at low values of *Excess comp*, indicated by a significantly negative coefficient on *High Bog*. We observe a significantly positive coefficient on *High Bog* in the high IO sub-sample (model 3), indicating that more sophisticated investors

are likely to be more sceptical of overpaid management as a result of low CD&A readability, irrespective of the magnitude of *Excess comp*.

We obtain qualitatively similar results in models (5) and (6), using the *Fog* measure of CD&A readability. In particular, the coefficient on *Excess comp*High Fog* is larger and more significant in model (6), where IO is low, and we observe evidence of obfuscation (negative coefficient on *High Fog*) when IO is low (model 6) but not when IO is high (model 5).

Taken together, our results corroborate, to some extent, the existence of a negative level effect of *High Bog (High Fog)* on *Dissent*, consistent with the obfuscation hypothesis and the results obtained by Hooghiemstra et al. (2017) in the UK context. However, we also observe a positive interaction effect of *High Bog (High Fog)* with *Excess comp*, indicating that the obfuscation effect of low readability gives way to heightened shareholder scepticism and dissent when CEO pay is highly excessive, consistent with the predicted effects of the processing fluency heuristic. This finding suggests that while obfuscation may be successful in attenuating shareholder dissent, this appears to be constrained to cases where CEO pay is at moderate levels. When CEO pay is clearly excessive, compared with firms sharing similar characteristics, unscrupulous firms looking to reduce the probability of shareholder dissent, would do better to disclose a 'plain English' (i.e., more readable) CD&A rather than attempt to obfuscate pay arrangements.

Whilst Hooghiemstra et al. (2017) find some evidence of increased dissent following low readability for firms with high levels of institutional ownership, we find that both the obfuscation (negative level effect) and processing fluency (positive moderating effect on *Excess comp*) effects of low readability are most pronounced for firms with more unsophisticated investors. This is consistent with the notion that unsophisticated investors are more likely to rely on heuristic processing. Consistent with Hooghiemstra et al. (2017), we find

no evidence of successful obfuscation for firms with high levels of institutional ownership, though we find evidence that low readability results in greater *Dissent* for high IO firms, as well as low IO firms.

4.3 Additional Analyses

4.3.1 Propensity Score Matching

Our methodology, so far, may be vulnerable to an endogeneity bias associated with omitted factors which may jointly determine CD&A readability and moderate the relationship between *Excess comp* and *Dissent*. It is plausible that when managers' interests are misaligned with those of shareholders, firms, in addition to awarding excessive CEO compensation, choose to disclose unreadable remuneration disclosures. In this case, our results may be driven by shareholders' aptitude to perceive underlying agency problems irrespective of CD&A readability. Thus, the heightened probability of SOP dissent for this group may not be affected by CD&A readability directly, and without accounting for omitted factors, our estimations could be biased.

In order to address the potential endogeneity problem, we conduct propensity score matching (PSM) based on the propensity for *Bog (Fog)* to be in the top quartile of its distribution (Q4), similar to Hooghiemstra et al. (2017). Specifically, we estimate a propensity score for *Q4 Bog* using a logit model containing the same covariates as our main model depicted in Eq. (3), except for *High Bog* and the *Excess comp*High Bog* interaction. We estimate the conditional probability of *Q4 Bog* using the logit model specified in Eq. (4) below.

$$\begin{aligned}
 \text{Q4 Bog}_{i,t} = & \alpha + \beta_1 \text{Excess comp}_{i,t} + \beta_4 \ln(\text{tenure}_{i,t}) + \beta_5 \text{IVOL}_{i,t} + \beta_6 \text{Neg_RET}_{i,t} \\
 & + \beta_7 \text{RET}_{i,t} + \beta_8 \text{Neg_RET}_{i,t} * \text{RET}_{i,t} + \beta_9 \text{Loss}_{i,t} + \beta_{10} \text{ROA}_{i,t} + \beta_{11} \text{Loss}_{i,t} * \text{ROA}_{i,t} \\
 & + \beta_{12} \text{EDec}_{i,t} + \beta_{13} \text{dROA}_{i,t} + \beta_{14} \text{EDec}_{i,t} * \text{ROA}_{i,t} + \beta_{15} \ln(\text{sales}_{i,t}) \\
 & + \gamma \cdot \text{year} + \delta \cdot \text{industry} + \varepsilon_{i,t}
 \end{aligned} \tag{4}$$

All variables are as defined previously (and in Appendix A). We use nearest-neighbour matching, with replacement, to match treatment observations (*Q4 Bog*) with control observations (where *Bog* is below median, i.e., $\leq Q2$). For each *Q4 Bog* observation, we first identify a pool of possible matches which (a) have below median levels of *Bog*, and also (b) share either a high or low (above or below median) level of institutional ownership (*%Inst*). We match with the observation within this pool having the closest propensity score, using a 0.01 calliper to ensure reasonable balance in propensity scores. We estimate the propensity scores and perform the matching using observations from the overpay sample only. We repeat the procedure above using *Fog* in place of *Bog*, thereby constructing a propensity-score matched sample for both readability measures.

Summary statistics for the matched sample are provided in Table 6. We analyse differences between treatment (*Bog* = Q4 or *Fog* = Q4) and control (*Bog* \leq Q2 or *Fog* \leq Q2) groups using standardised mean differences (SMD),²¹ two-sample t-tests, and nonparametric Wilcoxon matched-pairs signed-rank tests. Panel A of Table 6 presents results from matching on the *Bog Index*, while Panel B present results by matching on the *Fog Index*. In Panel A, we observe statistically significant differences in the outcome variable, *Dissent*, though the SMD are somewhat lower than 20, indicating that the differences are small. Thus, all else equal, treatment (*Q4 Bog*) observations are more likely to receive dissenting votes. In Panel B, we observe a similar difference in mean/median *Dissent* between treatment and control groups based on *Fog*.

²¹ Standardised mean differences (SMD) are calculated as the difference in means between treatment and control groups, divided by the common standard deviation. Cohen (1969) proposes SMD (sometimes referred to as *Cohen's d*) as a measure of effect size, suggesting a minimum value of 20 before acknowledging the existence of a small difference (with values 50 and 80 indicating medium and large differences, respectively). Thus we consider an SMD smaller than 20 to denote adequate balance in mean values between treatment and control groups.

We observe balance between treatment and control groups on all other observed variables in Panel A, according to the SMD and t-stat criterion, but a weakly significant difference in *IVOL* and $\ln(\text{Sales})$ according to the Wilcoxon signed-rank test. We observe a more significant difference in *IVOL* between groups in Panel B, as well as a weakly significant difference in *Neg_RET*. We consider therefore that the matching procedure achieves reasonable balance.

[Insert Table 6 about here]

We re-estimate Eq. (3) on the propensity score matched samples using conditional logistic regressions, and report the results based on the *Bog Index (Fog Index)* in Panel A (Panel B) of Table 7. We employ conditional logistic regressions in Table 7 in order to take into account the matched nature of the sample.²² In each case, we report results for the matched overpay sample as well as sub-samples based on high and low institutional ownership. The results are qualitatively consistent with those presented in Sections 4.1 and 4.2. In particular, we find that the coefficient on the *Excess comp*High Bog (Excess comp*High Fog)* interaction remains positive and significant in all regressions, following the PSM procedure. This strongly suggests that CD&A readability, itself, as opposed to underlying agency problems, moderates the link between excess CEO compensation and SOP dissent. We also find that the *Excess comp*High Bog (Excess comp*High Fog)* interaction in both Panels is larger (and more significant) for firms with low, compared with high, levels of institutional ownership, consistent with our findings in Section 4.2.

[Insert Table 7 about here]

4.3.2 *Shareholder Participation in the SOP*

²² Since the matching is performed on the basis of *High Bog (High Fog)*, this variable is omitted from the conditional logistic regressions.

While we examine, above, whether CD&A readability affects the voting preferences of shareholders taking part in the vote, a further important question is whether CD&A readability affects shareholders' decisions to vote at all. Hence, as an additional analysis, we examine whether the proportion of shares unambiguously voted during the SOP is affected by CD&A readability. When CD&A are excessively complex and costly to process, shareholders may simply decide not to vote at all. In this sense, low CD&A readability may exacerbate the problem of shareholder apathy, or disengagement in governance issues by (mainly small) shareholders.

To examine the possible effects on voting participation, we construct *SOP Participation* as the total number of shares voted either 'for' or 'against' in the SOP, divided by the number of shares outstanding at the date of the proxy statement. We therefore consider both 'abstain' votes and non-votes as non-participation, since no voting preference is expressed either way. We assume that the number of shares outstanding equates with the total number of eligible votes, as is the case with a one share, one vote policy. We therefore drop observations corresponding to firms with multiple share classes or unequal voting rights, which reduces our sample to 2,024 observations with ownership data.

We estimate an adaptation of Eq. (3) using OLS regression, with *SOP Participation* as the dependent variable, and including *%Inst* as an additional independent variable. The results are presented in Table 8. Consistent with the argument that small shareholders are less likely than institutional shareholders to engage with the SOP, we find that *SOP Participation* is significantly positively related with *%Inst*. For every additional 1% increase in *%Inst*, *SOP Participation* increases by approximately 0.34%. We find no evidence that *SOP Participation* increases with *Excess comp*, or CD&A readability, as the coefficients on *High Bog (High Fog)* and the *High Bog*Excess comp (High Fog*Excess comp)* interaction are insignificant. While we, therefore, conclude that variations in CD&A readability have no detectible marginal effect

on shareholder participation, preconceptions that CD&A and compensation arrangements are highly complex may still be a factor explaining disengagement by small shareholders.

[Insert Table 8 about here]

4.3.3 *Interplay with Negative Media Coverage*

While the primary focus of our paper is on the effects of CD&A readability on SOP voting decisions, we acknowledge that a complementary strand of literature suggests that negative or critical media coverage of related executives' pay may also increase SOP dissent (Hooghiemstra et al., 2015). Media scrutiny plays a role in highlighting cases of excess CEO compensation (Core et al., 2008), and firms receiving negative compensation-related media coverage tend to fare worse in the SOP (Hooghiemstra et al., 2015). If the way a CD&A is written affects the way shareholders perceive excess CEO pay, it may also affect the views formed by journalists. Thus, negative media coverage may be related with low CD&A readability, and may even amplify (or wholly explain) the effects of readability on Dissent.

The incidence of media articles scrutinising executive pay, however, has been shown to be related with Excess comp (Core et al., 2008), which is a strong predictor of *Dissent* (Brunarski et al, 2015). Therefore, an association between readability and media scrutiny (and ultimately, Dissent) may also exist if each is symptomatic of significant agency problems. On the other hand, if media scrutiny is related with agency problems (as suggested by Core et al., 2008), but CD&A readability is not, then we may observe no significant relationship between negative media coverage and readability.

We therefore construct and merge a database of negative media articles regarding executive compensation at companies in our sample. Following Hooghiemstra et al. (2015), media

articles relating to executive pay are obtained from the Lexis Nexis database using the following search string:

(Company Name) AND (CEO Name) w/20 (compensation OR salary OR bonus
OR option OR restricted stock)

We retain only articles published between the date of publication of the CD&A and the corresponding AGM where the SOP vote is held.²³ The remaining articles are categorised as ‘negative’ if at least one match is made with the search string for negative news developed by Core et al. (2008), outlined in Appendix C.

In Panel A of Table 9, we estimate a variation of Eq. (3) where *Neg_Cov* is substituted as the dependent variable. *Neg_Cov* is an indicator variable equal to 1 for observations where at least one negative article is published between the dates of the CD&A and the AGM, and 0 otherwise. The results indicate firm size (log of *Sales*) to be the most significant predictor of negative media coverage suggesting, as expected, the mainstream media are more likely to cover larger, more visible firms (Core et al., 2008; Fang & Peress, 2009). Consistent with Core et al. (2008), we also find *Excess comp* to be a significant predictor of negative media coverage in the overpay sample, with more excessively paid CEOs being significantly more likely to attract negative media coverage. The results in Panel A, however, provide no evidence of a relationship between CD&A readability and the propensity for negative media coverage. This lends some support to the notion that CD&A readability is not strongly related to agency problems, if we assume that media scrutiny is.

[Insert Table 9 about here]

²³ We obtain AGM dates from the EDGAR database. Specifically, AGM dates are listed as the ‘Period of Report’ on DEF 14A submissions. In a small number of cases where this rule did not appear to hold, we manually extract the AGM date from the text of the DEF 14A filing.

In Panel B of Table 9, we estimate a variation of Eq. (3) with *Dissent* as the dependent variable, however we include *Neg_Cov* as an additional independent variable, as well as a full set of interactions between *Neg_Cov*, *Excess comp*, and *High Bog (High Fog)*. The specification used in Panel B therefore tests the robustness of our findings when controlling for the publication of negative articles, and explores whether negative coverage amplifies/attenuates the effects of excess CEO pay and/or CD&A readability on *Dissent*. In general, we observe that our main inferences are robust when controlling for *Neg_Cov*, as the coefficient on *High Bog*Excess comp (High Fog*Excess comp)* is significantly positive in each case. We find that the inclusion of *Neg_Cov* has little impact, with insignificant coefficients estimated on all but one of the added covariates. This implies limited incremental predictability of *Dissent* from negative media coverage, above the original specification of Eq. (3).

4.3.4 Sensitivity to Length of Notice and the Volume of AGMs

Our main arguments regarding to the effects of CD&A readability on SOP voting outcomes rely on the presumption that shareholders have sufficient time to thoroughly review the CD&A. In practice, shareholders' attention is constrained (Hirshleifer & Teoh, 2003). Delayed reactions to news suggest it can take some time for investors to process new information (Hong et al., 2000), while extraneous events can be distracting, attenuating investor attention further (Hirshleifer et al., 2009). For example, Hirshleifer et al. (2009) report that post earnings announcement drifts are typically more pronounced when the volume of same-day earnings announcements is greater, as firms compete for investor attention.

Given the length of detail in the typical CD&A (approx. 10,000 words on average) and the volume of CD&As that are published (especially during the peak of the proxy season), shareholders may not always have the time or inclination to fully review each CD&A before voting. In particular, shareholders might be less likely to review the CD&A (and be affected

by its readability): (1) when the length of time between publication of the CD&A and the AGM is shorter; and (2) when there is a greater number of competing AGMs taking place at the same time, assuming that shareholders hold diversified portfolios.

We use these expectations to form the bases for additional sensitivity analyses, to further test whether our main results are likely to be driven by CD&A readability per se, as opposed to underlying agency problems that may simultaneously determine higher dissent and lower readability. We argue that if our results are driven by readability directly, then they will be attenuated when shareholders are more constrained in their ability to thoroughly review the CD&A. We therefore re-estimate Eq. (3) using sub-samples constructed by splitting on the variables *Short_notice* and *Busy*. *Short_notice* is an indicator variable equal to 1 when the number of days between publication of the CD&A and the AGM is below the median value, and 0 otherwise. *Busy* is an indicator equal to 1 if the number of AGMs of S&P 1500 firms occurring on the same day is above its median value, and 0 otherwise.

The results using the Bog Index (Fog Index) readability measure are presented in Panel A (Panel B) of Table 10. For brevity, we report only the results for the overpay sample. Overall, we observe that our main findings of a negative coefficient on *High Bog* (*High Fog*) and a positive coefficient on the *High Bog*Excess comp* (*High Fog*Excess comp*) interaction are strongest in the cases where there is a longer duration between CD&A publication and the AGM (*Short_notice* = 0), and when there are fewer competing AGMs taking place on the same day (*Busy* = 0). To reiterate, we expect that there is a greater likelihood of shareholders reviewing the CD&A under these circumstances, given lower constraints on investor attention. These findings therefore support the argument that the level of SOP dissent is affected by CD&A readability, directly.

[Insert Table 10 about here]

4.3.5 Using Continuous Readability and SOP Measures

We acknowledge that our methodology is somewhat limited since we employ both a binary measure of SOP dissent and a binary readability variable. While we adopt this methodological approach due to our expectation of non-linear effects as well as high skewness of SOP voting results, it is established that there are inherent problems with interpreting coefficients on interactions in logistic frameworks (Ai & Norton, 2003), as discussed in Section 4.1. We therefore present additional results in Table 11 from conventional OLS estimations of Eq. (3) but employing a continuous measure of SOP dissent (*% Dissent*) as well as continuous measures of CD&A readability, *Bog* and *Fog*. *% Dissent* is the untransformed percentage of shares voted ‘against’ in the SOP. *Bog* and *Fog* are the untransformed readability measures discussed in Section 3.3. We centre *Bog* and *Fog* to have a mean of zero, as recommended by Dawson (2014) in order to aid interpretation of the interaction effect.

[Insert Table 11 about here]

In general, our main result of a significantly positive coefficient on the *Bog*Excess comp* (*Fog*Excess comp*) interaction is robust to employing continuous variables in an OLS framework, with the exception of the coefficient on *Bog*Excess comp* in the overpay sample (though we do observe a significantly positive coefficient in the full sample). The weakening of our results within a linear framework suggests the existence of non-linear effects, as previously mentioned. We do not observe a significant coefficient on *Bog* or *Fog* in any of the models presented in Table 11, indicating our previous support of the obfuscation hypothesis (hypothesis 1) is limited. We also present in Table 11 the marginal effect of *Excess comp* at varying levels of *Bog/Fog*, namely at their 25th, 50th, and 75th percentile values, and observe that the marginal effect on *Excess comp* is significant yet increases with *Bog/Fog* in all cases. We express caution over interpretation of the results in Table 11, however, since they are likely

to be disproportionately driven by variations among the large number of observations with very low levels of SOP dissent.

4.3.6 *Potential Impact of Poor Performance on Readability*

Finally, as a further robustness test, we estimate a variation of Eq. (3) which includes interactions between the negative performance indicators *Neg_RET*, *Loss*, and *EDec*, and *High Bog (High Fog)*, as well as with *Excess comp*. To the extent that poor performance might be more difficult to explain, the estimated effect of CD&A readability on *Dissent* and the interaction with *Excess comp* might be driven by an omitted (but correlated) effect of particularly disappointing performance on shareholder dissatisfaction. The results are presented in Table 12, and indicate that our inferences are robust to this alternative explanation. In particular, the coefficients on the *Excess comp*High Bog (Excess comp*High Fog)* interaction remain significantly positive and are of a similar magnitude to those in Table 4. In addition, we observe a significantly positive coefficient on the *Excess comp*EDec* interaction across all models, suggesting that shareholders have even lower tolerance of high CEO pay when earnings have declined from the previous year.

[Insert Table 12 about here]

5. Conclusion

This paper provides evidence on the effect of remuneration disclosure readability on Say-on-Pay voting patterns in the US. Specifically, while low readability in the Compensation Discussion and Analysis (CD&A) may, in some cases successfully obfuscate excess CEO pay, resulting in lower SOP dissent (as reported by Hooghiemstra et al., 2017), this effect appears to be constrained to cases where Excess pay is low. At higher values of Excess pay (i.e., when CEO pay excesses are greater) we find that shareholders are *more* likely to vote ‘against’

abnormally high CEO compensation when the CD&A is written in a less readable manner. Thus, while at moderate CEO pay levels, a more difficult to read CD&A may ‘put off’ shareholders from challenging pay arrangements, low readability results in increased scepticism when CEO pay is highly excessive. Our findings are therefore consistent with Stathopoulos and Voulgaris’ (2016) view that many shareholders are put off from challenging executive pay arrangements unless pay levels are clearly excessive, but add that a Plain English CD&A seems to play an important role in quelling those concerns.

For S&P 1500 companies, we find that, although the probability of *Dissent* increases with the excessiveness of CEO compensation, this sensitivity is substantially reduced when a more readable CD&A is issued. For example, considering *Excess comp* of two standard deviations above the mean, the probability of *Dissent* is substantially higher for low readability firms than for high readability firms according to the Bog index measure (27.5% versus 17.8%), and is also over twice as high according the Fog index measure of readability (33.9% versus 13.8%). Our results are robust to possible confounding determinants of readability, and become muted in cases where shareholders are less likely to have sufficient time to review the CD&A. This suggests that CD&A readability directly affects the credibility of the ‘*compensation story*’ in cases where there is a clear need to justify the excessiveness of CEO pay.

We also show that the effects of CD&A readability on SOP dissent (both the level and moderating effects) are more pronounced when institutional ownership is low, and therefore shareholders in general are likely to be less sophisticated. This suggests that both the obfuscation effect of low readability, at low levels of Excess pay, and the processing fluency (heightened scepticism) effect of low readability, at higher levels of Excess pay, result from heuristic information processing by less-sophisticated investors. While Hooghiemstra et al. (2017) observe heterogeneity in the effects of readability on SOP dissent due to ownership structure (i.e., an investor sophistication effect), we provide an important contribution by

demonstrating that it is also strongly conditional on the degree of excessiveness in CEO pay. In particular, we show that the obfuscation effect of low readability appears to be limited to cases where *Excess pay* is moderate, and that it leads to increased dissent, even for firms with low institutional ownership, when CEO pay is highly excessive. By contrast, Hooghiemstra et al. (2017) suggest low readability increases scepticism and dissent only for firms with high levels of institutional ownership.

The findings in this paper therefore contribute to the literature by highlighting the role of remuneration disclosure readability in engendering shareholders' support of executive pay arrangements, and in particular when CEO pay otherwise appears highly excessive. Since shareholders' need for a justification of pay arrangements is likely to increase with the excessiveness of CEO pay, it is intuitive that the effect of Plain English remuneration disclosures (above less readable disclosures) in engendering trust and support of pay arrangements also increases in tow. Our results, therefore, also inform policymakers and practitioners regarding the intended consequences of Plain English language in increasing shareholders' trust in, and reliance on, corporate disclosures. However, they also highlight a possible unintended consequence, in that greater use of Plain English could potentially be used as a tool to lobby for greater shareholder support for excessively paid executives. Further research on potential heterogeneity in the effects of Plain English in other contexts (e.g., 10-K reports), and how these relate to investors' prior concerns and trust in management, may also shed important new light on the consequences of the Plain English initiative.

Our study is subject to number of limitations. Firstly, while the focus of our study is on the effects of readability on SOP outcomes, we do not comment on whether the *consequences* of SOP dissent are also affected. While we consider this to be outside of the scope of our study, it is likely to be fruitful avenue for future research, as it would help to identify whether or not the effect of Plain English on SOP outcomes advances shareholders' interests. Reduced

shareholder scepticism may be exploited by managers, thus resulting in even further pay excesses. On the other hand, increased trust between shareholders and managers may lead to improved governance arrangements.

Secondly, we examine only variations in the readability of CD&A, and not the content of CD&A, more broadly. It is likely that a host of other disclosure characteristics also affect shareholder approval of executive pay. For example, graphics and tables could be used to increase processing fluency or to highlight arguments that are favourable to management (and obfuscate others). Opportunistic choice of peers within comparator groups may also be used by managers to inflate support.

Finally, our approach implies that shareholders' prior concerns regarding pay arrangements (and consequently the effect of CD&A readability) are driven only by the results of benchmarking executive pay, however receptiveness to arguments in the remuneration disclosure may be primed by a host of other factors affecting shareholders' trust in management. For example, the strength of governance arrangements, or how management responded to previous shareholder votes. Whether shareholders' need for justification of excess pay (and therefore the potential impact of CD&A disclosure) is affected by factors outside of mere CEO pay levels, may therefore be a further potential avenue for future research.

Appendix A

Variable Names and Definitions

Variable	Definition
<i>Panel A: SOP voting and proxy variables</i>	
SOP Dissent (%)	The proportion of shares voting ‘against’ proposed executive pay in the Say-on-Pay advisory vote on executive remuneration.
Dissent	Indicator variable equal to 1 if greater than 30% of votes are cast ‘against’ proposed executive pay in the Say-on-Pay, and 0 otherwise. Dissent therefore takes the value 1 if less than 70% SOP support is obtained.
SOP Participation	The proportion of total shares outstanding voted either ‘for or ‘against’ proposed executive pay in the Say-on-Pay.
Bog	Bog Index value of Compensation Discussion & Analysis text.
High Bog	Indicator variable equal to 1 if Bog is above the median of all firm-years in the sample, and 0 otherwise.
Q4 Bog	Indicator variable equal to 1 if Bog is in the top quartile of all firm-years in the sample (i.e., above the 75% percentile), and 0 otherwise.
Fog	Fog Index value of Compensation Discussion & Analysis text.
High Fog	Indicator variable equal to 1 if Fog is above the median of all firm-years in the sample, and 0 otherwise.
Q4 Fog	Indicator variable equal to 1 if Fog is in the top quartile of all firm-years in the sample (i.e., above the 75% percentile), and 0 otherwise.
<i>Panel B: CEO-level variables</i>	
Total comp	The CEO’s total compensation at the fiscal year-end in \$000’s, as reported in the summary compensation table of the DEF 14A
Excess comp	Estimated measure of excess CEO compensation, defined in Section 3.2 as the residual from a predictive model of CEO Total comp. The estimated model is presented in Table 1.
Overpay	Indicator variable equal to 1 if Excess comp is positive, and 0 otherwise.
Underpay	Indicator variable equal to 1 if Excess comp is negative, and 0 otherwise.
Tenure	The tenure of the current CEO at the fiscal year-end, in years.
<i>Panel C: Firm-level variables</i>	
%Inst	Proportion of shares held by 13-F institutional investors at the end of the most recent quarter prior to the say on pay vote.
High IO	Indicator variable equal to 1 if %Inst is above the median of all firm-years in the sample, and 0 otherwise.
Low IO	Indicator variable equal to 1 if %Inst is below the median of all firm-years in the sample, and 0 otherwise.
TA	The value of the firm’s assets as at the fiscal year-end.
Sales	Firm sales as at the fiscal year-end.
Bk/Mkt	The book to market ratio calculated as (book value of assets)/(book value of liabilities + market value of equity).

S&P500	An indicator variable equal to 1 if the firm is a constituent of the S&P 500 index, and 0 otherwise.
IVOL	Idiosyncratic volatility of stock returns over 12 months to the fiscal year-end, calculated as the Root Mean Squared Error (RMSE) from a Fama-French 3 factor model.
ROA	Operating income scaled by total assets.
Loss	Indicator variable equal to 1 if ROA is negative, and 0 otherwise.
RET	Stock return over the 12 months to the fiscal year-end.
Neg_RET	Indicator variable equal to 1 if RET is negative, and 0 otherwise.
dROA	Percentage change in ROA between years $t-1$ and t .
EDec	Indicator variable equal to 1 if dROA is negative, and 0 otherwise.

Panel D: Contextual variables

Neg Cov	Indicator variable equal to 1 for observations where at least one negative article is published between the dates of the CD&A and the AGM, and 0 otherwise, as discussed in Section 4.3.3.
Short notice	Indicator variable equal to 1 when the number of days between publication of the CD&A and the AGM is below median, and 0 otherwise.
Busy	Indicator variable equal to 1 if the number of AGMs of S&P 1500 firms occurring on the same day is above median, and 0 otherwise.

Appendix B: Procedure for parsing the CD&A section of DEF 14A proxy statements.

1. We obtain the Complete Submission Text File (CSTF) for all DEF 14A filings relating to Annual General Meetings, and submitted by firms in the sample to EDGAR between 2010 and 2015 inclusive, using links published via the following url address:

`ftp://ftp.sec.gov/edgar/full-index/`

2. The ‘Compensation Discussion and Analysis’ section (CD&A) is isolated and extracted from each CSTF in the following manner:
 - a. DEF 14A section breaks are identified using HTML links between the text and the table of contents. Using HTML, a link to the CD&A section from the table of contents may take the following form:

`Compensation Discussion and Analysis`

With the beginning of the CD&A section specified similarly to the following:

`Compensation Discussion and Analysis`

All links from the table of contents (*href* attributes), as well as target section markers (*name* attributes) are extracted from each CSTF using regular expressions. A find-and-extract operation is performed using each of the two regular expressions below:

Regex1: `(href={0,1000}?)`

Regex2: `(name=[^>]+?>)`

Extracted links and targets within each CSTF are matched based on use of a common identifier, such that section markers (i.e., “*name=identifier*”>) are matched with the title given in the table of contents (i.e., the string between “*href=#identifier*” and the first subsequent “**”). Data from each CSTF are sorted according to the order in which targets occur in the text.

- b. The marker for the beginning of the CD&A section is identified based on keyword matching with titles from the table of contents. Specifically, references to CD&A sections are identified if the title contains each of the words “*compensation*”, “*discussion*”, and “*analysis*”, but none of the following words; “*report*”, “*vote*”, “*supplemental*”, “*appendix*”, and “*exhibit*”. The beginning of the CD&A was identified manually in 25 instances where it was found that this keyword search had failed to identify the appropriate marker.
- c. The end of the CD&A is determined as the position in the text where the section immediately after the CD&A begins. This poses some difficulty, however, as; (a) the CD&A is often divided into sub-sections, each of which is itemised and linked within the table of contents; (b) proxy statements do not follow a standard format, such that the ordering of sections differs between proxy statements; and (c) sections are often titled differently between proxy statements.

We therefore develop a procedure which screens each subsequent section title in turn to determine whether it is likely to be a sub-section of the CD&A. The beginning of the first subsequent section determined not to be a sub-section of the CD&A is taken to be the end of the CD&A. The process was developed inductively, with meticulous checking of the data at each stage to ensure the integrity of the procedure. The final form of the procedure is as follows:

- i. Remove all HTML code from extracted section titles.
- ii. Drop section markers where title consists only of a number, or white-space followed by a number.
- iii. Drop section markers where there is no title, or where the title consists only of white-space characters.
- iv. Where successive section markers are identified as being the beginning of the CD&A, keep only the first instance.
- v. Drop proxies where multiple non-successive CD&A section are identified, due to ambiguity.
- vi. Where greater than 12,500 characters of string exist between the beginning of the CD&A and the next section marker, this indicates that CD&A sub-sections are not linked, and the next section marker identifies the end of the CD&A.
- vii. Drop first section marker after CD&A if title contains either “report”, “vote”, “committee”, “capitalization”, “exelon”, “approval”, “compensation of executive officers”, or “investments”.
- viii. Drop section markers that occur fewer than 5,000 characters after the CD&A, as inspection of the data found that these relate to CD&A sub-sections.
- ix. For each remaining section marker subsequent to the CD&A in turn:²⁴
 - a) Identify as end of CD&A if title contains either “report”, “interlock”, “table”, or “insider”.
 - b) Otherwise, drop if contains either “guideline”, “highlight”, “philosophy”, “element”, “policies”, “policy”, “decision”, “vote”, “program”, “benefit”, “component”, “salary”, “analysis”, “target”, “incentive”, “process”, “principles”, “result”, “action”, “practice”, “performance”, “bonus”, “grants”, “data”, “summary”, “overview”, “introduction”, “business”, “2010 ”, “2011 ”, “2012 ”, “2013 ”, “2014 ”, “design”, “background”, “objectives”, “oversight”, “purpose”, “factors”, “named”, “governance”, “setting”, “role”, “NEO”, or “base salar”.
 - c) If is not yet dropped, identify as end of CD&A if title contains “risk”, or is greater than 12,500 characters after the CD&A and contains “committee”, [“executive” and “compensation”], or “officer”.²⁵ Otherwise drop the section marker and repeat step

²⁴ The keyword matches used below result from meticulous checking of each iteration to determine via visual inspection whether they exclusively indicate either (a) CD&A sub-sections, or (b) an appropriate ending point for the CD&A (i.e., the beginning of the section immediately after the CD&A).

²⁵ CD&A are commonly followed by the sections “Executive Compensation”, “Compensation of Executive Officers”, “Compensation Committee Report”. While these sections are also related to executive remuneration, they are generally very technical in nature, consisting mainly of tables and technical notes which support the more natural discussion in the CD&A. Since they may not therefore be amenable to textual analysis, we limit our focus to the CD&A.

- ix. for next successive section marker, until an appropriate ending is identified.
 - d. The text of each DEF 14A is split at the beginning and end of the CD&A, and only text between these positions are retained.
3. The text of CD&As are cleaned prior to analysis in the following manner:
- a. All white-space characters (e.g., newline “\n”, carriage return “\r”, and tab “\t” characters) are replaced with an ASCII non-breaking space character. Strings of multiple successive non-breaking spaces are replaced with a single non-breaking space.
 - b. All HTML character codes “ ”, “ ”, and “ ” are replaced with an ASCII non-breaking space.
 - c. All HTML character codes “’” and “'”, denoting an apostrophe, are removed from the text without leaving any space.
 - d. All instances of HTML character code “.”, denoting a period, are replaced with an ASCII period character.
 - e. All instances of HTML character code “&”, denoting an ampersand, are replaced with the word “and”.
 - f. Tables are identified in the text as string occurring in-between HTML tags “<table...>” and “</table>”. Similarly to Loughran & McDonald (2011),²⁶ tables are removed from the text if they contain greater than 15% numeric characters.
 - g. Each remaining segment of HTML code is replaced with a single ASCII non-breaking space character. HTML code segments are identified using regular expressions "<([^\>]+?)>" and "&([^\&]{1,7});".
 - h. Since HTML codes are replaced with spaces, a considerable amount of additional white-space is introduced. Therefore, any strings of multiple successive spaces are again replaced with a single space character.
 - i. Where string in-between alphabetic characters contains only a single hyphen “-” and non-breaking space characters, the space characters are removed. All other hyphens are replaced with a space character. These steps are important since end of line words are often broken by use of a hyphen. When ‘words’ are identified by matching strings against the master dictionary below, strings of alphabetic characters separated by a hyphen will be considered both (a) as a whole (removing the hyphen), and (b) separately (replacing the hyphen with a space).
 - j. Since sentences are identified by the occurrence of a period character “.” in the string, it is important to remove excess periods from the text which do not denote the end of a sentence (e.g., where periods are used as part of an

²⁶ While in Loughran and McDonald (2011, internet appendix) they remove tables which include greater than 25% numeric characters, in Loughran and McDonald (2016, p. 1221) they apply a threshold of 10%. Loughran and McDonald (2016) suggest that HTML table markers are less ambiguous in filings made after 2005.

abbreviation). The following steps are implemented to control for extraneous uses of the period character:

- i. Periods are removed from strings containing two or more periods, but no space characters (e.g., in the instance “w.r.t.”).
 - ii. Where the string between successive periods contains only non-breaking space characters, only a single period is retained.
 - iii. Where a period is immediately preceded by a space followed by a single alphabetic character, the period is removed.
 - iv. Periods between numeric characters are removed.
 - v. Periods are removed when immediately preceded by string consisting of a space followed by either “mr”, “inc”, “co”, “vs”, “ft”, “sq”, “jr”, “sr”, “ms”, “dr”, “no”, “ex”, “st”, “pp”, “I”, “II”, “III”, “IV”, “V”, “VI”, “VII”, “VIII”, “IX”, “X”, “XI”, “XII”, “XIII”, “XIV”, or “XV”.
 - vi. Where the string between successive periods does not contain alphabetic characters, the entire string is reduced to a single period.
- k. All apostrophes are removed (not replaced with a space).
 - l. All remaining non-alphabetic characters are replaced with a space character, while strings of successive spaces are again replaced with a single space.
 - m. The remaining CD&A text consists only of (a) alphabetic characters, (b) hyphens, (c) space characters, and (d) periods. Assuming that the text is cleaned appropriately, sentence breaks are identified by occurrences of a period character in the text (i.e., the number of period characters is equal to the number of sentences). Words are identified as string tokens of alphabetic characters (preceded/followed by a space or period), but which also match with any of the 84,667 words in the 2014 Loughran and McDonald Master Dictionary²⁷ with non-missing data on the number of syllables.²⁸ For hyphenated words, a match is first attempted on the string as a whole (excluding the hyphen). If this fails to match with a word, the hyphen is replaced with a space, and a match is attempted on each constituent string token. String tokens which fail to match with any word in the Master Dictionary are removed from the text.
 - n. Where the removal of string tokens results in successive periods where no words occur in the string between them, the extraneous periods are removed.

Finally, we ensure a standardised distribution of CD&A text, whereby each word and period are separated by a single non-breaking space character.

²⁷ Available via Bill McDonald’s website at: http://www3.nd.edu/~mcdonald/Word_Lists.html

²⁸ The Fog readability measure requires the ability to determine the number of syllables in each word.

Appendix C: Search string used to identify negative media articles

Media articles are coded as ‘negative’ if at least one match is found in the text or the article using the following search string:

(CEO name or CEO name’s or executive* or CEO*) w/25
(high* w/7 (salar* or bonus* or pay* or paid or compensat* or benefit*)
or excess* w/7 (salar* or bonus* or pay* or paid or compensat* or benefit* or option*)
or lofty w/7 (salar* or bonus* or pay* or paid or compensat* or benefit* or option*)
or hefty w/7 (salar* or bonus* or pay* or paid or compensat* or benefit* or option*)
or large w/7 (salar* or bonus* or pay* or paid or compensat* or benefit* or option*)
or rich w/7 (salar* or bonus* or pay* or paid or compensat* or benefit* or option*)
or big* w/7 (salar* or bonus* or pay* or paid or compensat* or benefit* or option*)
or outsize* w/7 (salar* or bonus* or pay* or paid or compensat* or benefit* or option*)
or huge w/7 (salar* or bonus* or pay* or paid or compensat* or benefit* or option*)
or generous w/7 (salar* or bonus* or pay* or paid or compensat* or benefit* or option*)
or exorbitant* w/7 (salar* or bonus* or pay* or paid or compensat* or benefit* or option*)
or fat* w/7 (salar* or bonus* or pay* or paid or compensat* or benefit* or option*)
or gargantuan w/7 (salar* or bonus* or pay* or paid or compensat* or benefit* or option*)
or bonanza* w/7 (salar* or bonus* or pay* or paid or compensat* or benefit* or option*)
or jumbo w/7 (salar* or bonus* or pay* or paid or compensat* or benefit* or option*)
or whopp* w/7 (salar* or bonus* or pay* or paid or compensat* or benefit* or option*)
or astound* w/7 (salar* or bonus* or pay* or paid or compensat* or benefit* or option*)
or ridiculous* w/7 (salar* or bonus* or pay* or paid or compensat* or benefit* or option*)
or stagger* w/7 (salar* or bonus* or pay* or paid or compensat* or benefit* or option*)
or handsome* w/7 (salar* or bonus* or pay* or paid or compensat* or benefit* or option*)
or lucrative w/7 (pay* or compensat* or option*)
or critic* w/7 (pay* or compensat*)
or best w/7 paid
or reap* w/7 million*
or self-serving
or largesse
or overpaid
or lavish
or perks
or perquisites
or windfall*
or earn* more than
or was paid more than
or receiv* more than
or made more than)

References

- Aggarwal, R., Saffi, P. A., & Sturgess, J. (2015). The role of institutional investors in voting: Evidence from the securities lending market. *The Journal of Finance*, *70*(5), 2309-2346.
- Ai, C., & Norton, E. C. (2003). Interaction terms in logit and probit models. *Economics Letters*, *80*(1), 123-129.
- Arnold, M.C., & Grasser, R. (2018). What is a fair amount of executive compensation? Outrage potential of two key stakeholder groups. *Journal of Business Finance & Accounting*, *45*(5-6), 651-685.
- Balsam, S., Boone, J., Liu, H., & Yin, J. (2016). The impact of say-on-pay on executive compensation. *Journal of Accounting and Public Policy*, *35*(2), 162-191.
- Bethel, J. E., & Gillan, S. L. (2002). The impact of the institutional and regulatory environment on shareholder voting. *Financial Management*, *31*(4), 29-54.
- Bloomfield, R. J. (2002). The “incomplete revelation hypothesis” and financial reporting. *Accounting Horizons*, *16*(3), 233-243.
- Bloomfield, R. J. (2008). Discussion of “annual report readability, current earnings, and earnings persistence”. *Journal of Accounting and Economics*, *45*(2-3), 248-252.
- Bonsall, S. B., Leone, A. J., Miller, B. P., & Rennekamp, K., (2017). A plain English measure of financial reporting readability. *Journal of Accounting and Economics*, *63*(2-3), 329-357.
- Bonsall, S. B., & Miller, B. P. (2017). The impact of narrative disclosure readability on bond ratings and the cost of debt. *Review of Accounting Studies*, *22*(2), 608-643.
- Brunarski, K. R., Campbell, T. C., & Harman, Y. S. (2015). Evidence on the outcome of say-on-pay votes: How managers, directors, and shareholders respond. *Journal of Corporate Finance*, *30*, 132-149.
- Bushee, B. J., & Miller, G. S. (2012). Investor relations, firm visibility, and investor following. *The Accounting Review*, *87*(3), 867-897.
- Bushee, B. J., & Noe, C. F. (2000). Corporate disclosure practices, institutional investors, and stock return volatility. *Journal of Accounting Research*, *38*, 171-202.
- Bushman, R. M., Gigler, F., & Indjejikian, R. J. (1996). A model of two-tiered financial reporting. *Journal of Accounting Research*, *34*, 51-74.
- Cai, J., & Walkling, R. A. (2011). Shareholders’ say on pay: Does it create value? *Journal of Financial and Quantitative Analysis*, *46*(2), 299-339.
- Cai, Y., Jiang, Y., Yin, H., & Zhang, H. (2017). Do managers choose the proxy material delivery method strategically? *Journal of Accounting, Auditing & Finance*, Forthcoming.
- Chen, X., Harford, J., & Li, K. (2007). Monitoring: Which institutions matter? *Journal of Financial Economics*, *86*(2), 279-305.
- Cohen, J. (1969). *Statistical power analyses for the behavioral sciences*, second ed. Academic Press: New York.

- Collins, D., Marquardt, B.B., & Niu, X. (2019). Equity-based incentives and shareholder say-on-pay. *Journal of Business Finance & Accounting*, 46(5-6), 739-761.
- Core, J. E., Guay, W., & Larcker, D. F. (2008). The power of the pen and executive compensation. *Journal of Financial Economics*, 88(1), 1-25.
- Correa, R., & Lel, U. (2016). Say on pay laws, executive compensation, pay slice, and firm valuation around the world. *Journal of Financial Economics*, 122(3), 500-520.
- Courtis, J. K. (1998). Annual report readability variability: tests of the obfuscation hypothesis. *Accounting, Auditing & Accountability Journal*, 11(4), 459-472.
- Cox, C. (2007). Speech by SEC chairman: Closing remarks to the second annual corporate governance summit. Retrieved from <https://www.sec.gov/news/speech/2007/spch032307cc.htm>
- Dawson, J. F. (2014). Moderation in management research: What, why, when, and how. *Journal of Business and Psychology*, 29(1), 1-19.
- De Jong, A., Huijgen, C. A., Marra, T. A., & Roosenboom, P. (2012). Why do firms go public? The role of the product market. *Journal of Business Finance & Accounting*, 39(1-2), 165-192.
- Del Guercio, D., Seery, L., & Woidtke, T. (2008). Do boards pay attention when institutional investor activists “just vote no”? *Journal of Financial Economics*, 90(1), 84-103.
- Dougal, C., Engelberg, J., Garcia, D., & Parsons, C. A. (2012). Journalists and the stock market. *The Review of Financial Studies*, 25(3), 639-679.
- El-Haj, M., Alves, P., Rayson, P., Walker, M., & Young, S. (2019a). Retrieving, classifying and analysing narrative commentary in unstructured (glossy) annual reports published as PDF files. *Accounting and Business Research*, Forthcoming.
- El-Haj, M., Rayson, P., Walker, M., Young, S., & Simaki, V. (2019b). In search of meaning: Lessons, resources and next steps for computational analysis of financial discourse. *Journal of Business Finance & Accounting*, 46(3-4), 265-306.
- Ernst & Young. (2018). *2018 proxy statements An overview of the requirements and observations about current practice*. Retrieved from [https://www.ey.com/publication/vwluassetsdld/2018proxystatements_06548-171us_30november2017/\\$file/2018proxystatements_06548-171us_30november2017.pdf?OpenElement](https://www.ey.com/publication/vwluassetsdld/2018proxystatements_06548-171us_30november2017/$file/2018proxystatements_06548-171us_30november2017.pdf?OpenElement)
- Ertimur, Y., Ferri, F., & Muslu, V. (2011). Shareholder activism and CEO pay. *The Review of Financial Studies*, 24(2), 535-592.
- Ertimur, Y., Ferri, F., & Oesch, D. (2013). Shareholder votes and proxy advisors: Evidence from say on pay. *Journal of Accounting Research*, 51(5), 951-996.
- Fang, L., & Peress, J. (2009). Media coverage and the cross-section of stock returns. *The Journal of Finance*, 64(5), 2023-2052.
- Frantz, P., Instefjord, N., & Walker, M. (2013). Executive compensation: A model of disclosure choice. *Journal of Business Finance & Accounting*, 40(9-10), 1184-1220.

- Greene, W. (2010). Testing hypotheses about interaction terms in nonlinear models. *Economics Letters*, 107(2), 291-296.
- Grossman, S. J., & Hart, O. D. (1980). Takeover bids, the free-rider problem, and the theory of the corporation. *The Bell Journal of Economics*, 11(1), 42-64.
- Hanley, K. W., & Hoberg, G. (2012). Litigation risk, strategic disclosure and the underpricing of initial public offerings. *Journal of Financial Economics*, 103(2), 235-254.
- Hartzell, J. C., & Starks, L. T. (2003). Institutional investors and executive compensation. *The Journal of Finance*, 58(6), 2351-2374.
- Hirshleifer, D., Lim, S. S., & Teoh, S. H. (2009). Driven to distraction: Extraneous events and underreaction to earnings news. *The Journal of Finance*, 64(5), 2289-2325.
- Hirshleifer, D., & Teoh, S. H. (2003). Limited attention, information disclosure, and financial reporting. *Journal of Accounting and Economics*, 36(1-3), 337-386.
- Hong, H., Lim, T., & Stein, J. C. (2000). Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies. *The Journal of Finance*, 55(1), 265-295.
- Hooghiemstra, R., Kuang, Y. F., & Qin, B. (2015). Say-on-pay votes: The role of the media. *European Accounting Review*, 24(4), 753-778.
- Hooghiemstra, R., Kuang, Y. F., & Qin, B. (2017). Does obfuscating excessive CEO pay work? The influence of remuneration report readability on say-on-pay votes. *Accounting and Business Research*, 47(6), 695-729.
- Hou, K., Karolyi, G.A., & Kho, B.C. (2011). What factors drive global stock returns?. *Review of Financial Studies*, 24(8), 2527-2574.
- Hsieh, C. C., Hui, K. W., & Zhang, Y. (2016). Analyst report readability and stock returns. *Journal of Business Finance & Accounting*, 43(1-2), 98-130.
- Hwang, B. H., & Kim, H. H. (2017). It pays to write well. *Journal of Financial Economics*, 124(2), 373-394.
- Ince, O. S., & Porter, R. B. (2006). Individual equity return data from Thomson datastream: Handle with care!. *Journal of Financial Research*, 29(4), 463-479.
- Khan, R., Dharwadkar, R., & Brandes, P. (2005). Institutional ownership and CEO compensation: a longitudinal examination. *Journal of Business Research*, 58(8), 1078-1088.
- Laksmana, I., Tietz, W., & Yang, Y. W. (2012). Compensation discussion and analysis (CD&A): Readability and management obfuscation. *Journal of Accounting and Public Policy*, 31(2), 185-203.
- Larcker, D. F., Scheider, R., Tayan, B., & Boyd, A. (2015). 2015 Investor survey: Deconstructing proxy statements – What matters to investors. Retrieved from <https://www.gsb.stanford.edu/faculty-research/publications/2015-investor-survey-deconstructing-proxy-statements-what-matters>
- Lawrence, A. (2013). Individual investors and financial disclosure. *Journal of Accounting and Economics*, 56(1), 130-147.

- Lee, Y. J. (2012). The effect of quarterly report readability on information efficiency of stock prices. *Contemporary Accounting Research*, 29(4), 1137-1170.
- Lehavy, R., Li, F., & Merkley, K. (2011). The effect of annual report readability on analyst following and the properties of their earnings forecasts. *The Accounting Review*, 86(3), 1087-1115.
- Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics*, 45(2-3), 221-247.
- Li, Y., & Zhang, L. (2015). Short selling pressure, stock price behavior, and management forecast precision: Evidence from a natural experiment. *Journal of Accounting Research*, 53(1), 79-117.
- Lo, K., Yang, S., & Zhang, J. L. (2014). Say-on-pay votes and compensation practices. Unpublished paper, University of British Columbia. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2377093
- Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, 66(1), 35-65.
- Loughran, T., & McDonald, B. (2014). Measuring readability in financial disclosures. *The Journal of Finance*, 69(4), 1643-1671.
- Loughran, T., & McDonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, 54(4), 1187-1230.
- Merkel-Davies, D. M., & Brennan, N. M. (2007). Discretionary disclosure strategies in corporate narratives: Incremental information or impression management?. *Journal of Accounting Literature*, 26, 116-196.
- Rennekamp, K. (2012). Processing fluency and investors' reactions to disclosure readability. *Journal of Accounting Research*, 50(5), 1319-1354.
- Robinson, J. R., Xue, Y., & Yu, Y. (2011). Determinants of disclosure noncompliance and the effect of the SEC review: Evidence from the 2006 mandated compensation disclosure regulations. *The Accounting Review*, 86(4), 1415-1444.
- Ross, I. M. (2013). The end of say on pay litigation? Retrieved from https://www.americanbar.org/content/dam/aba/administrative/young_lawyers/committee_newsletters/litigation_2013_fall_newsletter_sayonpay_jump.authcheckdam.pdf
- Schmidt, P.S., von Arx, U., Schrimpf, A., Wagner, A.F., & Ziegler, A. (2019). Common risk factors in international stock markets. *Financial Markets and Portfolio Management*, 33(3), 213-241.
- Securities and Exchange Commission (SEC). (1998). A Plain English handbook: How to create clear SEC disclosure documents. Retrieved from <https://www.sec.gov/pdf/handbook.pdf>
- Securities and Exchange Commission (SEC). (2006). Executive compensation and related person disclosure. Retrieved from <https://www.sec.gov/rules/final/2006/33-8732a.pdf>

- Semler Brossy. (2016). 2015 Russell 3000 say on pay results: End of year report. Retrieved from <http://www.semlebrossy.com/wp-content/uploads/SBCG-2015-Year-End-Say-on-Pay-Report.pdf>
- Serafin, A. T. (1998). Kicking the legalese habit: The SEC's Plain English disclosure proposal. *Loyola University Chicago Law Journal*, 29(3), 681-718.
- Shah, A. K., & Oppenheimer, D. M. (2007). Easy does it: The role of fluency in cue weighting. *Judgment and Decision Making*, 2(6), 371-379.
- Stathopoulos, K., & Voulgaris, G. (2016). The impact of investor horizon on say-on-pay voting. *British Journal of Management*, 27(4), 796-818.
- Strickland, D., Wiles, K. W., & Zenner, M. (1996). A requiem for the USA: Is small shareholder monitoring effective?. *Journal of Financial Economics*, 40(2), 319-338.
- Tan, H. T., Ying Wang, E., & Zhou, B. O. (2014). When the use of positive language backfires: The joint effect of tone, readability, and investor sophistication on earnings judgments. *Journal of Accounting Research*, 52(1), 273-302.
- Ulbricht, N. and Weiner, C., 2005. Worldscope meets Compustat: A comparison of financial databases. Unpublished paper, University of Mannheim and Humboldt University of Berlin. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=871169
- You, H., & Zhang, X. J. (2009). Financial reporting complexity and investor underreaction to 10-K information. *Review of Accounting Studies*, 14(4), 559-586.

Table 1
Cross-sectional regressions for estimating excess pay

Independent variables	Dependent variable
	Ln(Total comp) _t
Ln(tenure) _t	-0.024 (-0.96)
Ln(sales) _t	0.348*** (17.98)
S&P500 _t	0.394*** (8.49)
Bk/Mkt _t	-0.009 (-1.36)
ROA _t	-0.137 (-0.80)
ROA _{t-1}	-0.024 (-0.15)
RET _t	0.002*** (8.14)
RET _{t-1}	0.002*** (10.53)
Constant	3.233*** (10.85)
Observations	7,731
Year FE	YES
Industry FE	YES
Adj. R2	0.539

This table reports estimates from cross-sectional predictions of log total CEO compensation. We present OLS results using the full sample of 7,733 S&P 1500 firm-years between 2008 and 2014 for brevity, though calculate *Excess comp* in year t as the residual from estimating Eq. (1) using observations from years $t-2$ to t . All variables are defined in Appendix A. We include fixed effects for year and industry, following the Fama and French 12 industry classification.

T -statistics are shown in parentheses below coefficient estimates and are based on robust standard errors clustered at the firm level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively, based on a two-tailed test.

Table 2
Sample descriptive statistics

Variable	Mean	St. Dev.	1 st	Median	99 th
Bog	96.53	11.27	76	96	127
Fog	23.59	1.950	19.82	23.42	29.12
Total Comp (\$000's)	8,595	8,055	761.1	6,479	36,273
Excess comp	0.0763	0.551	-1.392	0.0749	1.560
SOP Dissent (%)	9.133	12.93	0.260	4.040	66.71
Tenure (years)	8.001	6.005	1	6.583	32
IVOL	1.425	0.691	0.623	1.259	3.622
RET	13.14	27.23	-65.01	13.84	80.22
ROA	0.092	0.085	-0.036	0.080	0.343
Loss	0.040	0.195	0	0	1
dROA	0.002	0.044	-0.122	0.002	0.138
Sales (\$millions)	11,850	33,760	105.0	2,788	138,074
TA (\$millions)	36,464	169,440	156.0	5,240	781,960
Mk/Bk	2.025	3.780	1.042	1.461	9.181

This table presents sample summary statistics for our full-sample of 2,686 S&P 1500 firm-year observations between 2010 and 2014. All variables are defined in Appendix A.

Table 3
Pearson correlation matrix

<i>Variable</i>		A	B	C	D	E	F	G	H	I	J
SOP Dissent (%)	A	1.00									
Ln(Total comp)	B	0.24*** (0.00)	1.00								
Excess comp	C	0.35*** (0.00)	0.53*** (0.00)	1.00							
Bog	D	0.04** (0.02)	0.03 (0.17)	0.09*** (0.00)	1.00						
Fog	E	0.06*** (0.00)	-0.08*** (0.00)	0.06*** (0.00)	0.73*** (0.00)	1.00					
Ln(Tenure)	F	0.12*** (0.00)	-0.03 (0.17)	0.13*** (0.00)	0.01 (0.57)	0.07*** (0.00)	1.00				
IVOL	G	0.09*** (0.00)	-0.32*** (0.00)	-0.01 (0.52)	0.05*** (0.01)	0.15*** (0.00)	0.04** (0.06)	1.00			
RET	H	-0.18*** (0.00)	0.03* (0.10)	-0.02 (0.28)	-0.01 (0.54)	-0.01 (0.55)	-0.03 (0.17)	-0.10*** (0.00)	1.00		
ROA	I	-0.10*** (0.00)	0.13*** (0.00)	0.03 (0.13)	-0.02 (0.23)	-0.02 (0.39)	-0.03* (0.09)	-0.07*** (0.00)	0.11*** (0.00)	1.00	
Ln(Sales)	J	0.02 (0.33)	0.73*** (0.00)	-0.14*** (0.00)	-0.04** (0.02)	-0.16*** (0.00)	-0.12*** (0.00)	-0.34*** (0.00)	-0.01 (0.67)	0.11*** (0.00)	1.00

This table presents Pearson correlations between the variables employed in our main analyses, for our full sample of 2,686 S&P 1500 firm-year observations between 2010 and 2014. All variables are defined in Appendix A. Two-tailed p-values are reported in parentheses below correlation coefficients.

Table 4
Determinants of high SOP dissent

Dependent variable: Dissent (1 if greater than 30% votes cast 'against')							
	Full sample		Underpay sample		Overpay sample		
Independent variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Excess comp	2.202*** (12.18)	1.909*** (8.31)	1.481** (2.18)	1.372 (1.56)	2.302*** (9.89)	1.891*** (6.45)	1.505*** (5.91)
High Bog		-0.134 (-0.62)		-0.065 (-0.11)		-0.304 (-1.00)	
Ex_comp X High Bog		0.543* (1.87)		0.246 (0.18)		0.738** (1.98)	
High Fog							-0.608** (-2.00)
Ex_comp X High Fog							1.513*** (4.14)
Ln(Tenure)	0.304** (2.41)	0.302** (2.39)	0.347 (1.00)	0.342 (0.98)	0.319** (2.42)	0.323** (2.43)	0.311** (2.44)
IVOL	0.278 (1.52)	0.274 (1.52)	0.443 (1.05)	0.447 (1.06)	0.233 (1.41)	0.229 (1.38)	0.236 (1.40)
Neg_RET	0.638*** (2.67)	0.620*** (2.60)	0.430 (0.83)	0.416 (0.84)	0.747*** (2.79)	0.722*** (2.71)	0.677** (2.54)
RET	-0.013** (-2.28)	-0.014** (-2.36)	-0.004 (-0.39)	-0.004 (-0.38)	-0.015** (-2.28)	-0.016** (-2.44)	-0.018*** (-2.65)
Neg_RET X RET	0.013 (1.51)	0.014 (1.51)	-0.013 (-0.77)	-0.014 (-0.82)	0.019* (1.89)	0.020* (1.95)	0.022** (2.13)
Loss	-0.685* (-1.65)	-0.734* (-1.67)	-1.097 (-1.12)	-1.106 (-1.14)	-0.643 (-1.38)	-0.719 (-1.42)	-0.744 (-1.55)
ROA	-5.730** (-2.06)	-5.635** (-2.00)	-12.206** (-2.18)	-12.181** (-2.16)	-4.546 (-1.55)	-4.421 (-1.48)	-4.802 (-1.58)
Loss X ROA	8.733 (1.58)	7.846 (1.50)	20.971 (0.83)	19.985 (0.85)	6.199 (0.93)	4.914 (0.72)	6.727 (1.00)
EDec	0.446*** (2.58)	0.434** (2.52)	-0.086 (-0.22)	-0.094 (-0.24)	0.545*** (2.72)	0.530*** (2.66)	0.531*** (2.62)
dROA	6.549** (1.98)	6.477* (1.94)	13.880*** (2.66)	13.836*** (2.65)	3.677 (1.06)	3.654 (1.04)	4.574 (1.23)
EDec X ROA	-7.499* (-1.74)	-7.574* (-1.75)	-18.61*** (-2.88)	-18.52*** (-2.89)	-3.583 (-0.76)	-3.837 (-0.81)	-4.031 (-0.79)
Ln(Sales)	0.228*** (3.13)	0.231*** (3.17)	0.379*** (3.03)	0.380*** (3.04)	0.179** (2.14)	0.182** (2.18)	0.176** (2.14)
Constant	-20.58*** (-15.42)	-22.54*** (-14.97)	-23.45*** (-8.97)	-22.19*** (-8.59)	-21.70*** (-14.18)	-21.09*** (-13.92)	-20.76*** (-13.99)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R ²	0.237	0.239	0.169	0.170	0.218	0.222	0.233
Observations	2,686	2,686	1,160	1,160	1,526	1,526	1,526

Marginal effect of Excess comp							
Average ME	0.133***	0.132***	0.037**	0.037**	0.197***	0.196***	0.197***
High Bog/Fog = 0		0.114***		0.036		0.160***	0.124***
High Bog/Fog = 1		0.146***		0.038		0.223***	0.252***

This table presents results for the determinants of high SOP dissent for our full sample of 2,686 S&P 1500 firm-years between 2010 and 2014, using a logistic regression framework. The dependent variable, *Dissent*, is an indicator variable equal to 1 if greater than 30% of SOP votes are cast ‘against’ proposed executive pay, and 0 otherwise. We also present results for the underpay/overpay sub-samples separately, defined as observations with negative/positive values of *Excess comp*, respectively. The independent variables are defined in Appendix A. We include fixed effects for year and industry following Fama and French’s 12 industry classification.

T-statistics are shown in parentheses below coefficient estimates and are based on robust standard errors clustered at the firm level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively, based on a two-tailed test.

We also present the average marginal effect of *Excess comp* over all observations in each regression, as well as the average marginal effect when *High Bog/High Fog* take the value 0 (1), separately, for models where they are employed.

Table 5
Ownership structure and determinants of high SOP dissent

Dependent variable: Dissent (1 if greater than 30% votes cast ‘against’)

	Full sample		Overpay sample			
	High IO	Low IO	High IO	Low IO	High IO	Low IO
Independent variables	(1)	(2)	(3)	(4)	(5)	(6)
Excess comp	2.203*** (5.86)	1.843*** (4.08)	2.422*** (5.31)	2.042*** (3.00)	1.891*** (4.98)	1.731** (2.12)
High Bog	0.501 (1.31)	-0.460 (-1.22)	0.819* (1.70)	-1.584** (-2.20)		
Excess comp X High Bog	0.345 (0.71)	0.830 (1.17)	0.079 (0.13)	2.305** (1.96)		
High Fog					0.152 (0.29)	-1.931*** (-2.94)
Excess comp X High Fog					1.143** (1.97)	2.691** (2.40)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R ²	0.315	0.319	0.273	0.426	0.282	0.433
Observations	1,070	1,071	660	549	660	549

This table presents sub-sample logistic regression results for the determinants of high SOP dissent, for 2,141 S&P 1500 firm-years between 2010 and 2014 where ownership data is obtained. We also present results based only on the overpay sample (observations with positive values for *Excess comp*). We present results for firm-years with High IO and Low IO (Institutional Ownership), separately, defined as above and below median values of *%Inst*, respectively. *%Inst* is the proportion of shares outstanding held by 13-F institutional investors (source: Thomson One Banker). All other variables are defined in Appendix A. The regressions also include the same set of controls as in Table 4, but we do not report the coefficients or the intercept in order to focus on the main variables of interest. We include fixed effects for year and industry following Fama and French’s 12 industry classification.

T-statistics are shown in parentheses below coefficient estimates and are based on robust standard errors clustered at the firm level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively, based on a two-tailed test.

Table 6
Comparison between propensity-score matched groups

Variables	Bog/Fog = Q4		Bog/Fog ≤ Q2		Analysis of differences		
	Mean	Median	Mean	Median	SMD	t-stat	Wilcoxon
<i>Panel A: Bog Index</i>							
Dissent	0.145	0.096	0.000	0.000	15.15	2.16**	2.18**
SOP Dissent (%)	12.570	11.121	4.865	5.547	10.07	1.44	0.37
Excess comp	0.480	0.483	0.385	0.385	-0.85	-0.12	0.03
Ln(Tenure)	1.984	1.929	2.048	2.015	7.60	1.08	0.49
IVOL	1.425	1.363	1.264	1.234	9.23	1.32	1.77*
Neg_RET	0.291	0.337	0.000	0.000	-10.08	-1.44	-1.45
RET	12.338	11.626	13.280	8.908	2.65	0.38	0.90
Loss	0.032	0.054	0.000	0.000	-10.92	-1.56	-1.57
ROA	0.090	0.095	0.080	0.071	-5.00	-0.71	-0.08
EDec	0.436	0.426	0.000	0.000	1.99	0.28	0.29
dROA	0.001	0.002	0.002	0.002	-0.92	-0.13	-0.76
Ln(Sales)	14.694	14.522	14.548	14.341	11.44	1.63	1.75*
<i>Panel B: Fog Index</i>							
Dissent	0.161	0.093	0.000	0.000	20.33	2.77***	1.42
SOP Dissent (%)	13.266	10.283	5.575	4.164	20.12	2.76***	1.97**
Excess comp	0.477	0.473	0.396	0.386	1.11	0.15	-1.16
Ln(Tenure)	2.019	2.001	2.079	2.069	2.60	0.36	-0.29
IVOL	1.490	1.337	1.308	1.239	25.80	3.53***	1.97**
Neg_RET	0.283	0.345	0.000	0.000	-13.42	-1.88*	-1.62
RET	12.249	11.698	12.014	7.993	2.05	0.29	0.82
Loss	0.035	0.050	0.000	0.000	-7.53	-1.07	-1.26
ROA	0.098	0.097	0.082	0.072	1.04	0.15	0.27
EDec	0.490	0.438	0.000	0.000	10.42	1.45	0.06
dROA	0.001	0.000	0.001	0.001	0.52	0.07	0.33
Ln(Sales)	14.627	14.613	14.412	14.510	0.97	0.14	1.65*

This table presents an analysis of differences between observations with top quartile values of our readability measures (Bog Index and Fog Index) and a propensity-score matched control group of firms with below median values (Bog/Fog ≤ Q2). In Panel A, we present estimates following matching based on Bog Index values of the CD&A, while in Panel B we present estimates following matching on the Fog Index.

We perform analyses of differences using Standardised Mean Differences (SMD), two-sample t-tests, and nonparametric Wilcoxon matched-pairs signed-rank tests. Results from these tests are presented in the last three columns. Standardised mean differences (SMD) are calculated as the difference in means between treatment and control groups, divided by the common standard deviation. Cohen (1969) suggests a minimum SMD value of 20 before acknowledging the existence of an economically small difference. For Wilcoxon and t-test results, *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 7
Results using the propensity-score matched sample: Determinants of SOP dissent

Dependent variable: Dissent (greater than 30% votes cast ‘against’)			
Independent variables	Full matched sample	High IO	Low IO
<i>Panel A: Bog Index</i>	(1)	(2)	(3)
Excess comp	1.380*** (2.80)	1.296* (1.70)	1.458 (1.07)
High Bog	<i>Omitted</i>	<i>Omitted</i>	<i>Omitted</i>
Excess comp X High Bog	1.371** (2.27)	1.743* (1.77)	4.527* (1.93)
Controls	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Pseudo R ²	0.248	0.287	0.589
Observations	812	429	287
<i>Panel B: Fog Index</i>	(4)	(5)	(6)
Excess comp	0.535 (1.23)	1.198* (1.79)	0.019 (0.01)
High Fog	<i>Omitted</i>	<i>Omitted</i>	<i>Omitted</i>
Excess comp X High Fog	2.287*** (3.88)	1.831** (2.08)	6.512*** (3.09)
Controls	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Pseudo R ²	0.259	0.353	0.582
Observations	814	431	282

This table presents conditional logistic regression results for the determinants of high SOP dissent, following a Propensity Score Matching (PSM) procedure. PSM is conducted for the overpay sample specifically, therefore the full matched sample contains only overpay observations. Results following matching based on *Bog Index* (*Fog Index*) are presented in *Panel A* (*Panel B*). Results are presented for the full propensity-score matched samples, as well as sub-samples based on High versus Low IO (institutional ownership). High (low) IO comprise firm-years with above (below) median values of %*Inst*.

All variables are defined in Appendix A. We include fixed effects for year and industry following Fama and French’s 12 industry classification.

T-statistics are shown in parentheses below coefficient estimates and are based on robust standard errors. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively, based on a two-tailed test.

Table 8
Regression on the proportion of shares voted

Dependent variable: SOP Participation (Proportion of shares voted)

Independent variables	(1)	(2)
%Inst (Institutional ownership)	0.337*** (5.26)	0.339*** (5.38)
Excess comp	0.970 (0.87)	2.719 (1.08)
High Bog	0.092 (0.24)	
Excess comp X High Bog	0.430 (0.53)	
High Fog		-1.631 (-1.23)
Excess comp X High Fog		-2.978 (-1.19)
Controls	Yes	Yes
Year FE	Yes	Yes
Industry FE	Yes	Yes
R ²	0.08	0.08
Observations	2,023	2,023

This table reports estimates from a cross-sectional OLS regression of the proportion of total shares outstanding voted either ‘for’ or ‘against’ in each issuers’ SOP, for 2,024 S&P 1500 firm-years between 2010 and 2014 where ownership data is obtained, and where a one-share one-vote system is used. All other variables are defined in Appendix A. We include fixed effects for year and industry following Fama and French’s 12 industry classification.

T-statistics are shown in parentheses below coefficient estimates and are based on robust standard errors clustered at the firm level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively, based on a two-tailed test.

Table 9
The role of negative media coverage

Independent variables	Full sample		Overpay sample	
<i>Panel A: Dependent variable = Neg_Cov</i>	(1)	(2)	(3)	(4)
Excess comp	0.497 (1.05)	0.619 (1.40)	1.139** (1.97)	1.117** (1.97)
High Bog	0.096 (0.44)		0.259 (0.55)	
Excess comp X High Bog	0.088 (0.17)		-0.412 (-0.60)	
High Fog		-0.152 (-0.64)		-0.107 (-0.22)
Excess comp X High Fog		-0.143 (-0.29)		-0.349 (-0.46)
Ln(Sales)	0.744*** (6.92)	0.743*** (6.83)	0.741*** (5.79)	0.749*** (5.83)
Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Pseudo R ²	0.155	0.155	0.194	0.196
Observations	2,686	2,686	1,526	1,526
<i>Panel B: Dependent variable = Dissent</i>	(5)	(6)	(7)	(8)
Neg_cov	-0.617 (-0.83)	-0.880 (-1.18)	-0.159 (-0.18)	-0.704 (-0.85)
Excess comp	1.835*** (7.47)	1.601*** (7.39)	1.813*** (5.47)	1.337*** (4.50)
Excess comp X Neg_cov	1.120 (1.22)	1.494 (1.58)	0.658 (0.71)	1.371 (1.46)
High Bog	-0.184 (-0.82)		-0.314 (-0.99)	
Excess comp X High Bog	0.648** (2.09)		0.803** (1.98)	
High Bog X Neg_cov	1.005 (1.01)		-0.228 (-0.19)	
Excess comp X High Bog X Neg_cov	-1.578 (-1.22)		-0.462 (-0.35)	
High Fog		-0.207 (-0.94)		-0.697** (-2.20)
Excess comp X High Fog		1.060*** (3.41)		1.708*** (4.24)
High Fog X Neg_cov		1.578 (1.61)		0.890 (0.68)
Excess comp X High Fog X Neg_cov		-2.258* (-1.78)		-1.871 (-1.31)
Controls	Yes	Yes	Yes	Yes

Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Pseudo R ²	0.240	0.247	0.222	0.234
Observations	2,686	2,686	1,526	1,526

This table presents results on the role of negative media coverage for our full sample of 2,686 S&P 1500 firm-years between 2010 and 2014. We also present results based only on the overpay sample (observations with positive values for *Excess comp*). Panel A presents results from logistic regressions where the dependent variable is *Neg_Cov*, an indicator variable equal to 1 for observations where at least one negative article is published between the dates of the CD&A and the AGM, and 0 otherwise. Panel B presents results from logistic regressions where the dependent variable is *Dissent*, but we include *Neg_Cov* and interactions between *Excess comp*, *High Bog (High Fog)* and *Neg_Cov* as additional independent variables. All variables are defined in Appendix A. We include fixed effects for year and industry following Fama and French's 12 industry classification.

T-statistics are shown in parentheses below coefficient estimates and are based on robust standard errors. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively, based on a two-tailed test.

Table 10
Robustness: Sensitivity to length of notice and volume of competing AGMs

Dependent variable: Dissent (greater than 30% votes cast 'against')				
	Short notice = 1	Short notice = 0	Busy = 1	Busy = 0
<i>Panel A: Regressions employing Bog Index</i>				
Excess comp	1.897*** (3.80)	1.935*** (4.67)	2.567*** (4.56)	1.354*** (3.53)
High Bog	0.161 (0.32)	-0.893** (-2.11)	0.111 (0.23)	-0.849* (-1.85)
Excess comp X High Bog	0.749 (1.11)	1.047* (1.85)	0.220 (0.32)	1.299** (2.46)
Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Pseudo R ²	0.260	0.255	0.242	0.254
Observations	596	930	831	695
<i>Panel B: Regressions employing Fog Index</i>				
Excess comp	1.774*** (4.19)	1.448*** (3.50)	1.993*** (4.98)	1.008** (2.53)
High Fog	-0.071 (-0.13)	-1.046** (-2.54)	-0.035 (-0.08)	-1.143** (-2.56)
Excess comp X High Fog	1.156* (1.70)	1.892*** (3.65)	1.240** (2.04)	1.949*** (3.71)
Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Pseudo R ²	0.265	0.265	0.257	0.264
Observations	596	930	831	695

This table presents sub-sample logistic regression results for the determinants of high SOP dissent for our overpay sample of 1,526 S&P 1500 firm-year observations between 2010 and 2014. Specifically, we split the sample by the dichotomous variables *Short Notice* and *Busy*, where; *Short Notice* takes the value 1 when the number of days between publication of the CD&A and the AGM is below median, and 0 otherwise; and, *Busy* takes the value 1 if the number of AGMs of S&P 1500 firms occurring on the same day is above median, and 0 otherwise. We present results employing the *Big Index (Fog Index)* readability measure in Panel A (Panel B).

All variables are defined in Appendix A. We include fixed effects for year and industry following Fama and French's 12 industry classification.

T-statistics are shown in parentheses below coefficient estimates and are based on robust standard errors. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively, based on a two-tailed test.

Table 11
Determinants of high SOP dissent – continuous measures

Dependent variable: SOP Dissent (%)				
Independent variables	Full sample		Overpay sample	
	(1)	(2)	(3)	(4)
Excess comp	8.773*** (11.99)	8.682*** (12.26)	15.977*** (12.11)	15.850*** (12.07)
Bog Index	0.011 (0.44)		-0.019 (-0.40)	
Excess comp X Bog Index	0.128** (2.36)		0.133 (1.24)	
Fog Index		0.151 (1.07)		-0.270 (-0.97)
Excess comp X Fog Index		0.877** (2.54)		1.345** (1.99)
Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Pseudo R ²	0.22	0.22	0.26	0.26
Observations	2,686	2,686	1,526	1,526
Marginal effect of Excess comp				
Bog/Fog at 25 th percentile	7.823***	7.559***	14.992***	14.127***
Bog/Fog at 50 th percentile	8.719***	8.545***	15.922***	15.639***
Bog/Fog at 75 th percentile	9.487***	9.687***	16.718***	17.390***

This table presents results for the determinants of high SOP dissent for our full sample of 2,686 S&P 1500 firm-years between 2010 and 2014, using an OLS regression framework. We also present results based only on the overpay sample (observations with positive values for *Excess comp*). The dependent variable, *SOP Dissent (%)*, is the untransformed proportion of shares voted ‘against’ proposed executive pay. *Bog Index* and *Fog Index* are the continuous readability measures, centred to have a mean of zero. All other variables are defined in Appendix A. We include fixed effects for year and industry following Fama and French’s 12 industry classification.

T-statistics are shown in parentheses below coefficient estimates and are based on robust standard errors clustered at the firm level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively, based on a two-tailed test.

We also present the marginal effect of *Excess comp* when *Bog Index/Fog Index* are at their 25th, 50th, and 75th percentile values.

Table 12
Interacting readability measures with poor performance

Dependent variable: Dissent (1 if greater than 30% votes cast ‘against’)

Independent variables	Full sample		Overpay sample	
	(1)	(2)	(3)	(4)
Excess comp	1.413*** (5.26)	1.389*** (5.82)	1.412*** (4.18)	1.222*** (4.09)
High Bog	0.122 (0.40)		-0.079 (-0.19)	
Excess comp X High Bog	0.607** (2.01)		0.881** (2.28)	
High Fog		0.126 (0.40)		-0.540 (-1.29)
Excess comp X High Fog		0.798*** 0.126		1.500*** -0.540
Neg_RET	0.795*** (2.58)	0.846*** (2.67)	1.106*** (2.65)	1.051*** (2.62)
Loss	-0.370 (-0.54)	-0.987 (-1.24)	-0.718 (-0.59)	-0.932 (-0.79)
EDec	0.111 (0.43)	0.147 (0.54)	-0.017 (-0.05)	-0.006 (-0.02)
Excess comp X Neg_RET	0.026 (0.07)	-0.056 (-0.16)	-0.392 (-0.84)	-0.580 (-1.20)
Excess comp X Loss	0.436 (0.49)	-0.073 (-0.11)	1.271 (1.08)	0.510 (0.48)
Excess comp X EDec	0.952*** (3.03)	0.848*** (2.73)	1.097*** (2.72)	0.934** (2.33)
High Bog X Neg_RET	-0.318 (-0.94)		-0.255 (-0.67)	
High Bog X Loss	-1.562 (-1.56)		-2.431*** (-2.78)	
High Bog X EDec	-0.170 (-0.52)		-0.204 (-0.57)	
High Fog X Neg_RET		-0.421 (-1.23)		-0.118 (-0.30)
High Fog X Loss		0.364 (0.37)		-0.511 (-0.49)
High Fog X EDec		-0.112 (-0.35)		0.008 (0.02)
Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Pseudo R ²	0.248	0.250	0.235	0.239
Observations	2,686	2,686	1,526	1,526

This table presents results for the determinants of high SOP dissent for our full sample of 2,686 S&P 1500 firm-years between 2010 and 2014, using a logistic regression framework. All variables are defined in Appendix A. We include fixed effects for year and industry following Fama and French’s 12 industry classification.

T-statistics are shown in parentheses below coefficient estimates and are based on robust standard errors clustered at the firm level. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively, based on a two-tailed test.