Are investor reactions shaping SEC filings writing?

Nihat Aktas, Eric de Bodt, and Can Deniz Dogan

This version: October 28, 2021

Abstract

This study examines whether information disclosed in Security Exchange Commission (SEC) filings complies with exogeneity conditions required to be used as independent variables in causal analyses. We use investor reactions to M&A announcements and related SEC filings as testbed. To this end, we rely on a large sample of hand collected SEC filings and consider two of their observable characteristics: (1) writing complexity, which is evaluated using textual analysis, and (2) time to file, which is identified thanks to availability of announcement and filing dates. Our regression analyses uncover a significant feedback effect from investor reactions to the writing complexity and timing of the SEC filings, a clear violation of strict exogeneity. These results are robust to many alternative empirical choices and confirmed using instrumental variables as identification strategy.

JEL classification: G34, G40

Keywords: SEC filings, endogeneity, mergers and acquisitions, investor reactions

---

Nihat Aktas is from WHU Otto Beisheim School of Management (nihat.aktas@whu.edu). Eric de Bodt is from Norwegian School of Economics (eric.debodt@nhh.no). Can Deniz Dogan is from WHU Otto Beisheim School of Management (can.dogan@whu.edu). We are grateful to Ettore Croci, Tingting Liu Christodoulos Louca and Burcin Yurtoglu for helpful comments and suggestions. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Are investor reactions shaping SEC filings writing?

The U.S. Securities and Exchange Commission (SEC) filings provide a wide range of information to investors and financial professionals on timely and regular bases. SEC filings beneficiate indeed from legal foundations that setup mandatory contents and delays to be respected. Moreover, these filings are easily accessible using the Electronic Data Gathering, Analysis, and Retrieval (EDGAR) system, freely available at https://www.sec.gov/edgar.shtml. Since more than fifteen years, the finance research community has seized this opportunity to collect new and detailed information to study firm behavior.

In the field of mergers and acquisitions (M&A), Boone and Mulherin (2007) use for the first time, to the best of our knowledge, SEC filings to provide an in-depth depiction of the M&A negotiation process. The authors uncover the importance of interactions between merging parties before the official announcement date (i.e., the private part of the negotiation process). The Boone and Mulherin (2007) contribution has since then triggered a wave of studies exploiting SEC filings content to examine determinants of corporate decisions (Boone and Mulherin, 2008; Officer et al., 2009, Aktas et al., 2010; Ahern and Sosyura, 2014; Masulis and Simsir, 2018; Liu and Officer, 2021, to list a few of them, limiting ourselves to M&As). These contributions use variables build thanks to information collected in SEC filings as determinants of interest in regression analyses. Causal interpretation of these results assumes that information collected in SEC filings is exogenous[1]: no latent factors correlated with the outcome of interest (e.g., abnormal returns, bid premium or mode of payment in the M&A case) are correlated with the content of the corresponding SEC filings. We investigate in this paper whether this assumption holds.

Is there any reasons to suspect that variables constructed with the help of SEC filings could violate the exogeneity assumption? We believe so. Information provided in SEC filings, even if framed in compulsory and strict regulations, is released by economic agents. These decision makers are optimizers, the main reasons of exogeneity violation in social sciences (Cunningham, 2021). The accounting literature, for example, abounds with such evidence. Already in 1993, Bartov confirms the presence of earning-smoothing practices around asset sales. Numerous follow up contributions confirm the existence of such practices and study their determinants and consequences (see, e.g., Bergstresser et al., 2006; Erickson et al., 2006; Crocker and Slemrod, 2007; Strobl, 2013; Caskey and Laux, 2017). Whether or not these practices reflect fraudulent behaviors, they challenge the exogeneity condition.

[1] By exogeneity, we refer to the classic strict exogeneity condition required (but not necessarily sufficient) for causal interpretation of regression analysis results.

Exogeneity violation by SEC filings-based independent variables has potentially far-reaching implications for research in finance with the generalization of algorithmic textual analysis. The internet has completely revolutionized access to information. In the pre-internet age, access to public information gathered by official agencies was difficult and extremely time-consuming, imposing to consult physically archived documents. Under the impetus of internet, these public archives have been digitalized and put on-line. The SEC EDGAR database is just one such example, the list of U.S. governmental agencies' databases available on-line being endless, but an impressive one. As of June 2021, EDGAR provides access to 14,380,919 filings across 706 form types, starting in 1994 (https://www.sec.gov/dera/data/dera_edgarfilingcounts). The exploration of the content of such amount of information has led to the intensive use of automated text processing algorithms in accounting and finance research. Following Boone and Mulherin (2007), the use of these algorithms has targeted M&A related SEC filings but, nowadays, many different SEC filings and other sources of textual contents (such as earnings call transcripts and company press releases) are concerned: corporate culture (Li et al., 2021), CEO behavior (Aktas et al., 2016), financial constraints (Bodnaruk et al., 2015; Hoberg and Maksimovic, 2015; Buehlmaier and Whited, 2018), innovation (Mukherjee et al., 2017), product-market competition (Hoberg and Phillips, 2010 and 2016), cybersecurity risk (Florackis et al., 2020), to list a few. If information disclosed in SEC filings is (to some point) endogenous, one may legitimately worry about causal interpretation of these empirical results.

Testing whether SEC filings' content complies with exogeneity requirements is by definition challenging because economic agents' private information and their specific incentives are not observable to the econometrician. But firm decisions leave traces that can be tracked. Our empirical strategy rests on testing whether observable SEC filing characteristics are correlated with investor reactions to M&A announcements. The existence of such correlation would indeed indicate that the content of the corresponding SEC filings depends on factors that drive investor reactions at the deal announcement. Such feedback effect should not exist under strict exogeneity. If moreover this feedback effect appears to have a causal taste, suspicions of endogenous information disclosure will be reinforced. Indeed, this will provide indications that investor reactions per se have a direct impact on how and when decision-makers choose to release information.

We focus on SEC filings related to M&A announcements for two main reasons. First, M&A deals are characterized by a specific announcement date, even if determining the exact announcement date is sometimes empirically challenging (Mulherin and Simsir, 2015). The existence of an announcement date allows to measure investor reactions relying on the well-established event study method. Second, filings to the SEC occur after the announcement of the merger agreement, allowing us to examine the feedback effect from investor reactions to the writing complexity of the corresponding SEC filing.

The first characteristic of filed SEC forms that we scrutinize is writing complexity. SEC filings writing respond indeed to strict criteria, especially since the inception of the Plain Writing Act of 2010. Quoting the SEC itself (https://www.sec.gov/plainwriting.shtml), "The SEC, like other federal agencies, must write documents in plain writing, defined under the Act as writing that is clear, concise, well-organized, and follows other best practices appropriate to the subject or field or audience". Already in 1998, with the publication of the handbook "A Plain English Handbook: How to Create Clear SEC Disclosure Documents" (https://www.sec.gov/pdf/handbook.pdf), the SEC showed its commitment to communicate with investors in easily accessible language. Under The Plain Writing Act of 2010, SEC filing writing should therefore be highly standardized and uncorrelated with deal characteristics. A correlation with investor reactions at the deal announcement, if present, will signal some departure with this body of writing principles.

To quantity writing complexity, we use the Fog index, introduced in 1952 by Gunning (Gunning, 1952), originally to evaluate grade school reading material.[2] The Fog index has been widely used to quantify writing complexity in prior literature. Li (2008) is a representative example of its use in the recent period. The author quantifies the readability of U.S. firm annual reports and test the relation between readability and earnings (that appear to be positive, higher readability being correlated with higher earnings). As stressed in Loughran and McDonald (2020), the Fog index is easy to compute and to tabulate. Moreover, its interpretation (the number of years of formal education needed to understand the document in a first reading) is appealing. Despite limitations pointed out by the authors, the Fog index wide use in accounting and finance literature is a key benefit in our case because it allows us to test for the presence of feedback effects from investor reactions to document complexity, as measured in existing empirical works.

Merging firms have also significant leeway in delays to file M&A related forms into EDGAR. For example, in case of merger, form S-4 must be filed no later than twenty business days before the meeting during which security holders will vote on the merger proposal. But it is the decision of the merging parties to call for the security holders meeting. Therefore, in addition to analyzing the writing complexity of filed SEC forms, testing whether firms time their filings makes sense. Two questions are of interest here: do firms account for investor reactions to decide when to file and does this decision affect significantly the time to closing of the transaction? The first complements the writing

[2] We test the robustness of our results using seven alternative measures of writing complexity borrowed from prior literature.

complexity-based test of exogeneity and the second potentially provides indications of strategic information disclosure behavior to impact the time to completion.

Our baseline analysis relies on forms S-4 and DEFM14A filed by acquirers (as identified using their SEC Central Index Key – CIK - code) but we replicate our tests on the S-4 and DEFM14A forms filed by targets as a robustness check. SEC filings are available in EDGAR from 1994 and onwards. We therefore collect M&A transactions in the Refinitiv SDC database (SDC) from 1994 up to end of 2019. Focusing on completed control transactions between public U.S. acquirers and targets with deal value of at least one million US, we obtain a sample of 4,633 transactions. Next, we collect S-4 and DEFM14A forms in EDGAR and apply several filters to keep only filings allowing us to compute the Fog index thereafter (e.g., we exclude filings with less than fifty words in the background section). In particular, we screen manually all filings to exclude those with reduced or missing background, reason or opinion sections. We set also manually beginning and end of section markers and computed the number of words in each of them. A last filter is the availability of necessary information in the Center for Research in Security Prices (CRSP) database to compute acquirer cumulative abnormal returns (CAR) over a 3-day window centered on the transaction announcement date. This procedure leaves us with 1,698 transactions, 1,658 S-4 forms and 40 DEFM14A forms.

We study the correlation between SEC filings writing complexity and time to file by regressing the Fog index and the number of days between the announcement date and the filing date on acquirer CAR at the announcement date. We add a large list of control variables (the number of filings by deal, tender offer, unsolicited offer and full cash dummy variables, the toehold if any, the merging parties relative size, the acquirer size, Tobin's Q and leverage, the target firm size, Tobin's Q and leverage). We do not have an *a priori* on whether these variables are potential determinants of the writing complexity or the time to file (because, to the best of our knowledge, we are the first to run such regressions) but these variables are classically used in acquirer CAR regressions. They are therefore potentially correlated with our independent variable of interest. Controlling for their effect helps to isolate the direct relations between investor reactions, writing complexity and time to file and therefore, to test for the presence of a feedback effect. We report many variations around this baseline approach: different subsets of control variables to fight the classic bad control issue (Angrist and Pischke, 2009), models with interaction variables in an attempt to identify potential factors driving our main results, robustness checks with alternative econometric estimators and measures of writing complexity, and adoption of an instrumental variables approach to further alleviate omitted variable bias, which is the main source of endogeneity in the present case. These tests confirm our baseline results.

Our baseline regression delivers two unambiguous messages: acquirer CAR are associated (i) negatively with SEC form writing complexity and (ii) positively with the time to file, with both relations being strongly statistically significant: following more negative investor reactions around the transaction announcement, acquirers file relatively more complex SEC forms, and they file faster to the SEC. The simple fact that the null hypothesis of no relation is rejected is consistent with the violation of the exogeneity assumption in SEC filing information disclosure. We also observe that the writing complexity of SEC filings is positively correlated with tender offers, acquirer and target size, and target leverage. In complement, we observe that the time to file is shorter in case of tender offers and when SEC forms are filed by larger acquirers. Unsolicited transactions are also filed faster. Using the Amihud Illiquidity (2002) ratio and the acquirer industry average CAR as instruments for the acquirer CAR, we obtain comparable results. This confirms that our results are unlikely to be driven by unobservables.

In an attempt to gain a better understanding of the sources of the considered feedback effect, we investigate the role of the product-market competition using the Hoberg and Phillips (2010) similarity score, the corporate governance using CEO duality, the CEO power using CEO pay-slice as in Bebchuck et al. (2011) and CEO overconfidence using the longholder proxy (Malmendier and Tate, 2005). None of these variables appear to mediate the relation between writing complexity, the time to file and investor reactions. Assuming that this absence of results is not the by-product of type two error, SEC filings information disclosure endogeneity appears to be a pervasive phenomenon that does not hang upon firm characteristics.

We report several additional results. Noteworthy, replicating our baseline specification but using the time from transaction announcement to closing or the time from filing to closing reveals that only the time to file is related to investor reactions. This result is consistent with acquirers timing strategically the filing of the SEC forms after negative market reactions to speed up the completion of the deal.

Our results call for caution when interpreting regression results using variables based on information collected in SEC filings. The existence of a feedback effect from investor reactions to two main SEC filing attributes indicates that SEC filing-based information release is not absent of endogeneity presumption. This should not come out as a surprise. The accounting literature report since many years numerous evidence of comparable results concerning earnings disclosure (Bartov, 1993; Bergstresser et al., 2006; Erickson et al., 2006; Crocker and Slemrod, 2007; Strobl, 2013; Caskey and Laux, 2017). Potential consequences of our findings are however far reaching in the light of the regular use of SEC filings information and automated textual analysis in modern empirical corporate finance (Boone and Mulherin, 2007; Hoberg and Phillips, 2010; Bodnaruk et al., 2015; Hoberg and

Maksimovic, 2015; Aktas et al., 2016; Hoberg and Phillips, 2016; Buehlmaier and Whited, 2017; Mukherjee et al., 2017; Florackis et al., 2020; Li et al., 2021, among many others).

We review the literature related to our work in Section 1. Next, we summarize relevant SEC filings regulations and EDGAR features in Section 2. Section 3 is dedicated to the presentation of our empirical design. In Section 4, we report our results, before concluding.

1. Literature review

We start this review with the case of earnings disclosure as a typical example of information released by optimizing agents. It is important to establish that information distortions do exist in such cases as this is the problematic that we address. Next, we highlight the importance of SEC filings as a source of information in modern empirical corporate finance, especially with the generalization of automated textual analysis technics. This motivates our investigations. Finally, we review indicators used to evaluate writing complexity, a key variable in our analyses.

1.1. Information disclosure behaviors

Private information is valuable and therefore, economic agents incorporate value effects of information release when deciding to disclose it. These strategic behaviors have been observed and analyzed in many different contexts in accounting and finance, from insider trading to M&A negotiations. We mainly focus on earnings announcements because, like for SEC filings, there are strict accounting regulations in-place that prohibit earnings manipulation to fool investors. We also provide evidence of strategic behaviors in the context of M&A-related information disclosure because we will use M&As as our empirical testbed.

Bartov (1993) studies whether managers smooth earnings through the timing of assets salesUsing data collected in Compustat over the 1987 to 1989 period, Bartov confirms the existence of such practices. Bergstresser et al. (2006) focus on assumptions used by managers to value pension assets and their implication on disclosed earnings. Using data collected in Compustat and Internal Revenue Service (IRS) form 5500 filings, the authors gather 12,719 observations on 2,442 firms over the period. 1991 to 2002. Their conclusions are unambiguous: "firms use higher assumed rates of return when they prepare to acquire other firms, when they are near critical earnings thresholds, and when their managers exercise stock options", a set of evidence strongly supporting deviations from accounting rules regulating earnings computation. Erickson et al. (2006) investigate fifty cases of firms accused by the SEC of accounting fraud during the period 1996 to 2003 to test whether executive equity incentives are associated with these illegal behaviors. Interestingly (because in contrast to common believes), the authors did not find such relation: executive equity-based compensation seems not to drive accounting fraudulent practices. Crocker and Slemrod (2007) go one step further and show that some degree of earnings management tolerance is required to design managerial contracts that maximize profits, assuming earnings constitute private information observed by managers. This conclusion invites us to avoid assimilating too fast endogenous information release choices with fraudulent manipulations. Strobl (2013) explore the relation between earnings manipulation and the firm's cost of capital. The author shows that earnings manipulation is correlated with the business cycle, depending on firms' earnings profile. This dependence explains that earnings manipulation can influence the firm's cost of capital. Let us finally mention Caskey and Laux (2017) who study the relation between board governance and managers' incentives to select information disclosed accounting reports. The authors conclude that effective reporting oversight is required to limit such managers' leeway.

This short journey in the accounting earnings disclosure literature leaves no place to doubt of the existence of endogenous information disclosure by optimizing economic agents, a main source of exogeneity violation in social sciences (Cunningham, 2021). Earnings disclosure is only one such example in accounting and finance. In the context of M&As, there is also evidence of the existence of such behaviors, even if research in this area is relatively limited in comparison to the accounting literature. Ahern and Sosyura (2014) uncover that bidders in stock mergers appear to generate more press releases to influence their stock price after the initiation of merger negotiations. Imperatore et al. (2021) examine whether targets choose peers willingly for valuation purposes in the fairness opinion section of the corresponding SEC filing. In particular, the authors document that targets employ lower-valued peers when they anticipate higher litigation risk, which is the case when the target CEO is retained *ex post* in the combined firm. In a related research and relying also on SEC filings, Eaton et al. (2021) provide evidence consistent with target-firm advisors selecting strategically peers with high valuation multiples to negotiate higher takeover prices.

1.2. SEC Filings data collection and algorithmic textual analysis in corporate finance

Boone and Mulherin (2007) initiate a new stream of contributions that investigate in depth the M&A negotiation process and its relation with M&A outcomes. The authors' starting point is the apparent low level of competition in the M&A market (the average number of bidders reported in the SDC database typically fluctuates around 1.1, as for example in Liu and Officer (2021)). Boone and Mulherin very cleverly undertake to collect by hands M&A related SEC filings 14A, S-4 (for mergers) and 14D (for tender offers) to study the background section describing the M&A negotiation process. Decomposing this process into a private and a public part (Hansen, 2001), the authors report that, for a sample of four hundreds takeovers over the period 1989 to 1999, on average 9.49 potential buyers are contacted by the selling firm and its investment banks, 3.75 engage in a confidentiality/standstill agreement and 1.29 submit a private written offer. Moreover, one out of two of these transactions can be characterized as (formal or informal) auctions. Going into the information content of SEC filings provide a new and completely different picture of the competition among bidders in the M&A market. Numerous contributions using M&A related SEC filings will follow. Boone and Mulherin (2008) use the same source of information to characterize the depth of competition among bidders prior to the public announcement of the transaction and reject the presence of a winner's curse in M&A auctions. Officer et al. (2009) study the relation between target information asymmetry and acquirer returns in the case of acquisitions of private companies. The authors collect information on target's financial data in acquirer SEC filings 8-K, S-4, S-1 and S-3. Results indicate that acquirer returns are significantly higher when using stock as method of payment to acquire difficult-to-value targets. Aktas et al. (2010) collect information on the M&A selling procedure in SEC filing to study the relation between latent competition and bid premium in negotiated transactions. Results obtained using a sample of 1,774 M&A transactions over the 1994 to 2007 period (927 auctions and 847 negotiations) confirm that, even in one-on-one negotiations, latent competition is at play and drive-up prices. Masulis and Simsir (2018), using again information collected in M&A SEC filings, report that target-initiated transactions are common (35.4% in a sample of 1,268 transactions over the 1997 to 2012 period), motivated by a combination of economic weakness, financial constraints, and negative economy-wide shocks. These transactions lead to lower deal premia. Liu and Officer (2021) use SEC filings to collect bid revisions during the private and public parts of the M&A process. In the analyzed sample of 1,324 M&A transactions over the 1994 to 2016 period, 74.6% display a positive price revision during the private part, while the authors find only 8.8% such revisions during the public part. Additional analyses lead to conclude that target managers pursue shareholder wealth maximization.

On-line access to EDGAR from all over the world, thanks to the development of internet coupled with exponential growth of computation power and storage space of modern computers, has moreover spurred the use of SEC filings-based information in all fields of corporate finance. Hundreds of thousands of filings can easily be downloaded using FTP (file transfer protocol) and automatically processed using algorithms tuned to this task (the main difficulty being the loosely structured organization of filings that complexifies parsing SEC forms in their respective sections and sub-sections). Emblematic of these trends, Hoberg and Phillips (2010, 2016) introduce similarity scores as a measure of product market competition. These are obtained by collecting product descriptions in SEC filings 10-K, coding the used vocabulary in vectors of binary variables and computing distances between these vectors. The use of product market similarity scores sheds new light on competition in the M&A market and product differentiation strategies. Other sources of textual information are subject to intensive mining, such as earnings call transcripts and press releases. Li et al. (2021) collect 209,480 earnings call transcripts and rest on machine learning algorithm (the word embedding model) to characterize corporate culture. The authors show that corporate culture correlates with business outcomes such as operational efficiency, risk-taking, earnings management, and executive compensation design. Aktas et al. (2016) compute the number of occurrences of singular personal pronoun ("I" and its variations) relative to plural personal pronoun ("we" and its variations) to characterize CEO personality (narcissism in reference to existing works in psychology) and to study whether CEO personality traits affect M&A outcomes. This appears to be the case: narcistic CEOs complete negotiations faster, among others. Bodnaruk et al. (2015) parse SEC filings 10-K using a lexicon specific to financial constraints vocabulary. The authors show that their text-based measure is weakly correlated with measures classically used in the financial literature and are better predictors of liquidity events (such as dividend omissions or increases, equity recycling or underfunded pensions). Hoberg and Maksimovic (2015) also use SEC filings 10-K to obtain measures of financial constraints, adding the distinction between equity and debt financing ones. The reported results put forward the significant differences in characteristics of firms facing these two types of constraints. In the same vein, Buehlmaier and Whited (2018) build measures that discriminate between access to equity markets, debt markets and external financial market in general. In all cases, constrained firms earn higher returns, and especially so for debt constrained firms. Mukherjee et al. (2017) question the relation between corporate taxes and innovation. The authors complement classic patents and R&D based measures of innovation with new product introduction proxies obtained thanks to textual analyses of 98,221 unique firm press releases and controls for endogeneity thanks to the use of staggered changes in state-level corporate tax rates. They conclude that corporate taxes hinder innovation by reducing incentives and discouraging risk-taking. Florackis et al. (2020) assess firms' exposure to cyber-attacks. To build their measure of cybersecurity risk, the authors extract the discussion on risk factors from SEC filings 10-K, use a sub-sample of firms subject to major cyber-attacks to collect a dictionary of words representative of these risks, and used the resulting dictionary to predict firm exposure to new attacks, implementing an approach similar to Hoberg and Phillips (2010) similarity scores. Firms more exposed to cybersecurity risk earn higher expected returns on average.

The use of these new textual analysis algorithms has undoubtedly contributed to enrich quantitative data sources used in empirical corporate finance. However, giving a causal taste to the observed correlations could be hazardous if it appears that information disclosures by optimizing economic agents is exposed to endogeneity bias. Our endeavor is to investigate whether it is the case for information collected in SEC filings.

1.3. Writing complexity

Our primary measure of writing complexity is the Fog index, introduced by Gunning in 1952. The index evaluates the readability of English writing and estimates the number of years of formal education needed to understand a text on first reading. This index has been widely used since its introduction: a search on "Fog index" in Google Scholar generates 8.160 results on 7/27/2021, and Li (2008), an application of the Fog index to evaluate annual report complexity, collect 1,976 citations on its own. This is certainly to be credited to its attractive features: the Fog index is easy to compute and to tabulate, and it leads to an intuitive interpretation.

Loughran and McDonald (2011) show however that financial texts have specificities and that word lists developed in other disciplines can misclassify word tones. A similar line of reasoning lead Loughran and McDonalds (2014) to question the suitability of the Fog index to characterize complexity of financial document writings: complex words in common vocabulary are not necessarily complex when used in financial documents and the number of words in a sentence may reveal itself hard to measure when dealing with documents containing itemized lists, headings and numerous abbreviations. The authors argue in favor of using the file size as an alternative. Despite of these limitations, Loughran and McDonalds (2020) acknowledge that the Fog index remains widely used in accounting and finance. A typical application of the Fog index to SEC filings 10-K is Baxamusa et al. (2018). The authors study a sample of 1,581 strategic alliances and report that when the partner annual reports readability is worse, investor reactions to the announcement of the alliance (CAR) are lower, especially when investors suspect insufficient due diligence before the alliance's formation or when the partners are from different industries. In the eyes of investors, opaque annual reports increase the probability of mismatch.

The wide use of Fog index motivates our choice of this measure as main outcome of interest. This enables us to test whether information disclosure in SEC filings is related to investor reactions. It is important to emphasize that, as we study the determinants of the Fog index heterogeneity in the cross-section of firms, limitations put forward in Loughran and McDonald (2014) are not hampering the interpretation of our results. Moreover, we complement results obtained with the Fog index with seven alternative measures of writing complexity in our robustness checks.

2. Legal context

We introduce in this section elements of U.S. legislations that are relevant to our research question. We start by describing forms that must be filed to the SEC in case of M&A transactions and by introducing EDGAR, the main tool put at disposal of the public to access them. Next, we introduce the Plain Writing Act of 2010, that allows us to posit our main null hypothesis.

2.1. M&A SEC filings and EDGAR

The issuance of securities publicly traded on a U.S. security exchanges is governed by the requirements of the U.S. Securities Act of 1933. Firms issuing such securities are commonly referred as U.S. public firms. These are required to file a set of financial reports and other information to the SEC under the U.S. Securities Exchange Act of 1934. Mergers are primarily regulated by the SEC through proxy rules pursuant to Regulation 14A of this later. The two main forms in case of statutory mergers (also referred as one-step merger[3]) are form DEFM14A and form S-4.[4] These two filing types follow comparable structures, including a sophisticated background, reasons for the merger, and opinion of financial advisors sections.

Form DEFM14A, filed under the U.S. Securities Exchange Act of 1934, is the merger proxy statement. This form, which is filed by the target to inform its shareholders, includes information about the company, the acquirer and the merger transaction itself. The document typically provides details on the transaction background (a description of the negotiations and discussions between the parties and their representatives), the transaction terms (a description of the merger and the approvals that must be obtained by the parties), the target board recommendation (including the motivations), a summary of opinions obtained by the target about the fairness of the transaction, a set of financial statements and projections related to the merger, information about the shareholder meeting and voting procedures and information regarding directors, officers and beneficial owners of five percent or more the target's shares. .

If the consideration in the merger includes acquirer securities, the acquirer will be required to file a registration statement on form S-4 under the U.S. Securities act of 1933. This form complements information disclosed by the target in its proxy statement with information on the acquirer and acquirer securities, including financial projections. In such case, most often the parties combine their respective disclosure documents into a single document (simply called the proxy statement) filled on form S-4. Once filed, the SEC will notify within ten calendar days if it intends to review the filing and, if so, the SEC has up to thirty calendar days to provide comments to the parties. When definitive, the

[3] In one-step merger, the acquirer negotiates a definitive merger agreement with the target, which typically must first be approved and declared advisable by the target's board of directors.

[4] In case of tender offer or exchange offer directly addressed by the acquirer to the target's shareholders (known as two-step acquisition), the acquirer files a form TO-T that contains information about the acquirer and the terms and conditions of the offer. If the acquirer issues securities as full or partial consideration, it must also files a form S-4. For these transactions, the target's board of directors must also file a form 14D-9 containing a recommendation statement.

form S-4 is sent to the target's shareholders, accompanied by a proxy card required to vote, typically at least twenty to thirty business days before the target's shareholders meeting.

The SEC gives freely access to SEC filings on-line thanks to EDGAR, described as "the primary system for compagnies and others submitting documents under the Securities Act of 1933, the Securities Exchange Act of 1934, the Trust Indenture Act of 1939 and the Investment Company Act of 1940" (https://www.sec.gov/edgar/about). EDGAR figures are impressive: 3,000 filings are processed by day and 3,000 terabytes of data are served to the public annually. As of March 2001, EDGAR contains 13,199,753 filings across 695 different form types. Among these, 14,175 are S-4 forms and 5,255 DEFM14A forms. Figure 1 displays the evolution of the number of DEFM14A filings from 1993 to 2019 (for this illustration, we limit ourselves to form DEFM14A because S-4 forms are not necessarily related to M&A transactions). The classic aggregate M&A wave pattern (Harford, 2005) is clearly apparent, with the peak in the nineties before the internet bubble burst, the rebound of the M&A market in 2005, the sharp drop during the 2008 financial crisis and revival thereafter.

2.2. From the Plain Writing Act to testable hypotheses

The Plain Writing Act, signed in 2010 by President Obama, requires that federal agencies use so called plain-writing in every document that the agency issues. The concept of plain writing has a long history that can be traced back to the sixteenth century, meaning written in clear and straightforward language (see https://en.wikipedia.org/wiki/Plain_English). The Act imposes to federal agencies a set of initiatives to improve documents writing such as training their employees to this form of writing, implementing a process for overseeing compliance with plain writing requirements, informing the public that agency complies with the Act on agency's website, and designating points-of-contact for the public to interact with the agency on the implementation of the Act. Deadlines to implement the Act were in July and October 2011 (see https://www.plainlanguage.gov/law/).

In 1998, long before the Plain Writing Act, the SEC already published a handbook explaining how to create clear SEC disclosure documents, accessible at https://www.sec.gov/pdf/handbook.pdf. In the SEC own words, "This handbook shows how you can use well-established techniques for writing in plain English to create clearer and more informative disclosure documents". Composed of twelve chapters, the handbook explains what is a plain English document and provides guidelines on document organization, writing and readability evaluation. Among the common issues hampering readability (and to be therefore avoided), the handbook list long sentences, passive voice, weak verbs, superfluous words, legal and financial jargon, numerous defined terms, abstract words, unnecessary details and unreadable design and layout. The handbook goes beyond just writing recommendations, addressing also information to be disclosed. Questions such "Does the document highlight information that is important to investors?" or "Is any important information missing?" are explicitly raised (see pages 11 and 12).

SEC forms writing is clearly subject to a strict body of rules and this leads us to formulate our main null hypothesis: DEFM14A and S-4 filings' writing complexity should be uncorrelated with M&A transactions features and, in particular, with investor reactions at the transaction announcement. Uncovering a statistically significant correlation would indicate a feedback effect from investor reactions to SEC form writing, a violation of exogeneity for variables computed from information collected in DEFM14A and S-4 filings when used in regression analyses of M&A outcomes.

In addition to this main hypothesis, the leeway from which the merging parties beneficiate to effectively file forms DEFM14A and S-4 into EDGAR (see Section 2.1) suggests that the time to file is also worthwhile to be scrutinized. A statistically significant correlation between investor reactions at the deal announcement and the time to file would indeed be additional evidence of the existence of a strategic component in merging parties filing behavior.

3. Method

3.1. Empirical strategy

Ideally, having access to private information sets of the merging parties, we could verify whether information disclosed in the SEC forms faithfully informs investors, independently of M&A outcomes anticipated by these later. In such case, testing for exogeneity of information disclosed in SEC filings would reduce to some simple test of correlation between investor anticipations and the information content of these filings. Private information sets of the merging parties is however not observable.

Previous contributions uncover the existence of feedback loops from investors to corporate decision makers in the case of M&A transactions. In particular, Luo (2005) and Kau et al. (2008) show that CEOs take into account investor reactions to M&A announcements to decide whether or not to pursue their acquisition attempts. The potential existence of such feedback effects motivate our empirical strategy that relies on testing whether SEC filings writing complexity is correlated with investor reactions to M&A announcements. Under the Plain Writing Act of 2010 (and SEC recommendations in place from 1998), this should not be the case, as explained in Section 2. Moreover, using instruments, we are in a position to check the robustness of our results to the presence of omitted variables and to verify whether the (potential) correlation between SEC filings writing complexity and investor reactions is the side effect of some latent factors.

The case of M&As provide us an interesting setup for such explorations because these transactions represent major events in corporate life, with financial stakes at play being most often very important. This potentially motivates economic agents to take into account value effects of information disclosure. Another interesting aspect of the M&A setting is that the announcement date is well identified, even if collecting it might be challenging in some cases (Mulherin and Simsir, 2015). In addition, SEC forms to be filed are clearly defined and their reporting occurs well after the announcement of the M&A agreement, opening the doors to feedback effect analyses.

Merging parties benefit finally from some leeway to decide when to file SEC forms into EDGAR (see Section 2). We therefore analyze whether the time to file also appears to be conditioned by investor reactions as additional investigations.

3.2. Sample

To collect a representative sample of DEFM14A and S-4 forms filed to the SEC, we start by extracting from SDC M&A transactions from 1994 to 2019 that comply with the following criteria: both the acquirer and the target must be U.S. public firms, the transaction must be completed, the deal value must be at least one million US$, the four-week bid premium must be reported and the acquirer must own more than fifty percent of the target after the transactions (in 98.36 percent of the transactions present in our final sample, the acquirer fully acquires the target). We exclude buybacks, recapitalizations, exchange offers, acquisitions of partial interest and acquisitions of remaining interest. By doing so, we obtain a sample of 4,633 transactions, a figure comparable to M&A sample size reported in previous recent contributions taking into account the filters that we impose (see, e.g., Liu and Officer, 2021).

For this sample of M&A transactions, we gather DEFM14A and S-4 filings. Starting from the list of all DEFM14A and S-4 filings in EDGAR between 1994 and 2019 (19,430 filings), we first require that the filing dates fall between the announcement and closing dates of an M&A transaction present in our sample. Next, we associate SEC filings to M&A transactions. To this end, we proceed in three steps: we start by matching acquirer and target CIK unique identifiers[5]; next, we search for perfect overlap of the company names; finally, we match manually residual cases looking for correspondence in company names. This name matching procedure leads to a sample of 4,841 filings for 3,530 M&A transactions present in our sample. Note however that the filings' entry in the EDGAR provides information on the filing company, which is either the acquirer or the target, but the entry does not mention both merging parties. Our identifier and name matching procedure is therefore exposed to false positive (finding matches based on one name that correspond to another M&A transaction). We

[5] EDGAR uses CIK identifiers. We build a CIK to GVKEY (company identifier in Compustat) and CUSIP (company identifier in SDC) correspondence table using linking tables provided by Compustat on Wharton Research Data Services (WRDS).

check therefore manually our matches and exclude these cases.[6] This leads us to identify 3,480 filings, allocated to 2,684 deals.

All S-4 and DEFM14A forms are not suited to compute the Fog index, our measure of writing complexity. We check that the filings contain enough writing material by manually extracting sections related to the background of the deal, reasons for the merger and opinion of financial advisors and drop filings in which one of these sections contains less than fifty words. 3,049 filings, concerning 2,465 M&A transactions, survive this criteria.

A last filter is the availability of necessary information in the Center for Research in Security Prices (CRSP) database to compute acquirer cumulative abnormal returns (CAR) over a 3-day window centered on the transaction announcement date. This procedure leaves us with 2,569 filings, bearing on 2,224 deals. Focusing on the SEC forms filed by acquirers and keeping the first filing per deal to avoid overlapping filings for the same M&A transaction, we are left with 1,698 transactions, 1,658 S-4 forms and 40 DEFM14A forms. These are the filings that compose our baseline sample.

Table 1 reports the number of M&A transactions by year in column 1 and the corresponding aggregate transaction value in US$ million (transaction values are collected in SDC) in column 2. The classic aggregate M&A wave pattern is clearly apparent (Harford, 2005), paralleling the wave pattern in the yearly number of DEFM14A filings reported in Figure 1. In transaction value, 1998 is the peak year in our sample, even if large transactions clearly occur also during the last decade as in Alexandridis et al. (2017).

We replicate the procedure described hereabove to collect SEC forms filed by the target, using the target CIK as identifier. We obtain 849 DEFM14A and 22 S-4 forms, for a total of 871 M&A transactions. This sample of target side filings will be used to check the robustness of our results on the acquirer side.

3.3. Variables

Writing complexity and time to file

Our main dependent variable is the Fog index, introduced in Gunning (1952). This writing complexity measure was designed to evaluate the number of years of formal education needed to understand a text on first reading. It is computed as:

$$Fog_i = 0.4 \times \left(\frac{\#\, words_i}{\#\, sentences_i}\right) + 100 \times \left(\frac{\#\, complex\ words_i}{\#\, words_i}\right) \tag{1}$$

[6] To identify false positive matches, we search for the company name of the other involved party in the filing, and require that the name appears at least once. Company names in the EDGAR and SDC databases are however sometimes too detailed and/or contain abbreviations that differ between the two data sources. Manual search is important to solve many such matching intricacies between filings and M&A transactions.

where i is the document subscript, Fog stands for Fog index, $\# \ words$ is the number of words in the document, $\# \ sentence$ is the number of sentences, and $\# \ complex \ words$ is the number of complex words. Complex words are defined as words consisting of three or more syllables. Using the Master Dictionary of Loughran and McDonald (https://sraf.nd.edu/textual-analysis/resources), we identify 48,061 words in the dictionary, which consist of three or more syllables and search for them in the filings. When processing DEFM14A and S-4 forms to evaluate the Fog index, we start by removing the header, exhibits and tables, which we identify by their HTML (hypertext markup language) tags. Additionally, we delete all HTML tags themselves from the documents because these are formatting commands that do not affect the writing complexity in itself. Using the cleaned documents, we identify sentences by searching for dots, question marks and exclamation marks.

We report in Table 2 the arithmetic average (column 2), standard deviation (column 3) and first, second and third quartiles of the Fog index distribution in our sample of 1,698 filings. The arithmetic average and median are close to each other, indicating that strong outliers are unlikely and that the distribution is mostly symmetric (the skewness coefficient is -0.23, unreported). Reading DEFM14A and S-4 filings apparently requires around 22 years of formal education, but this interpretation must be taken with caution because of the intrinsic limitations of the Fog index for accounting and financial documents (see Section 2). We report also in column 3 of Table 1 the evolution of the average Fog index through years. While stable during the first decade, a clear positive trend appears from 2003 and onwards. By the end of the period, the average Fog index has increased by close to twenty percent. Apparently, SEC filings writing complexity got worse despite the implementation of the Plain Writing Act of 2010. Both in terms of sample average and trend through time, the distribution of the Fog index in our sample of M&A-related SEC filings is comparable to the one reported in prior literature focusing on 10-K filings (see, e.g., Dyer et al., 2017).

Our second dependent variables of interest is the time to file, simply computed as the number of days between the M&A announcement date (as reported in SDC) and the DEFM14A or S-4 filing date. The average time to file is 66 days in our sample as reported in Table 2. To put that number in context, this is forty one percent of the average time to close the transaction (i.e., the number of days between the announcement date and the closing date). Completing SEC form to be filed is clearly a demanding task and firms (and their investment bankers and lawyers) dedicate time to it. We report in column 4 of Table 1 the time-series of the average time to file. No time trend is discernable, with the average time to file fluctuating between 40 days (in year 2001) and 129 days (year 1994) throughout our sample period.[7] We report in Table 3 correlation coefficients between our main variables of interest

[7] The average time to file in 1994 appears to be abnormally high relative to the sample average of 67 days. This is due to the National Energy/Alexander Energy transaction with a time to file of 607 days, and a time to close of 639 days. Note that excluding this single observation does not affect our findings (unreported).

and their corresponding level of statistical significance. The correlation between the Fog index and the time to file is negative and statistically significant at the 1% level: filings featured by lower readability are filed faster, an intuitive result if we are willing to consider that complying with the Plain Writing Act takes time.

We also use other variables on the left-hand side of regression specifications in various robustness checks and additional analyses. For a better flow of the discussion, these variables will be introduced in the course of the results' presentation.

Investor reactions

We use the classic event study approach introduced in Fama et al. (1969) to measure investor reactions at the transaction announcement date. Following Fuller et al. (2002), we select the so-called beta-one model as return generating process, because it does not require the estimation of risk factors using an estimation window that can be contaminated by other events. Moreover, Brown and Warner (1985) show that short term event studies are highly robust to the choice of the return generating model. Abnormal returns (AR) are therefore simply obtained as:

$$AR_{it} = r_{it} - r_{mt}, \tag{2}$$

where subscript i stands for the M&A transaction and t for time, r_{it} is the return on firm i, and r_{mt} is the equal-weighted market index return. We select a three-day event window centered around the transaction announcement date reported in SDC.[8] Cumulative abnormal returns (CAR) are obtained by summation:

$$CAR_i = \sum_{t=-1}^{+1} AR_{it} \tag{3}$$

CAR descriptive statistics are provided in Table 2. Three days average acquirer CAR are -2.3 percent, as typically obtained in samples of large M&A transactions targeting public firms (see Betton et al., 2008). Around the filing dates however, we observe no discernable average investor reactions but first and third quartile values (respectively -2.1 percent and 1.7 percent) indicate a large degree of heterogeneity across transactions. Maybe more importantly with respect to our research question, we see in Table 3 that acquirer CAR at the announcement date are negatively correlated with writing complexity (statistically significant at the 10% level), a first indication that investor reactions might affect SEC form writing. Also worthwhile to be noted, the correlation between acquirer CAR and the time to file is negative and highly significant at the 1% level, another indication that investor reactions might affect firms filing decisions. Note finally that the absolute value of acquirer CAR is positively

[8] As a robustness check, we also extend the event window to five days to better control for anticipations and potential errors in announcement dates reported in SDC. The results are similar with this alternative event window (unreported). Moreover, our main findings are also robust to the use of the classical market model as a normal return generating process instead of the beta-one model (unreported).

correlated with the file size, an intuitive result: M&A transactions that generate stronger investor anticipations' revisions are most probably more complex and associated with longer filings.

Control variables

In the absence of existing references on the determinants of SEC filings writing complexity, we select control variables known to be correlated with acquirer CAR, our main independent variables of interest (for a recent reference, see Alexandridis et al., 2017). This is our best chance to limit the risk to be contaminated by endogenous omitted variables (and moreover this contributes to improve the power of our tests to the extent that these control variables contribute enough to the reduction of the residuals variance).

The list of controls variables considered in our baseline multivariate analyses is long. In the full regression specification, we account for the number of filings collected by deals (even if we keep in the sample only the first one), the structure of the deal (tender offer dummy), the nature of relation between the merging parties (unsolicited dummy), the mode of payment (cash dummy), the presence of toehold, the relative size between the merging parties, bidder and target characteristics (size, Tobin's Q and leverage ratio), and whether the target has also submitted a separate filing for the same transaction (joint filing dummy). All variables are defined in Appendix A and descriptive statistics are provided in Table 2. We limit ourselves to note that our M&A sample includes very few full cash transactions (3.5%). This is because filing S-4 form to the SEC is not compulsory for the acquirers in full cash transactions.

Table 2 reports descriptive statistics on many other variables that are used in robustness checks and additional analyses: textual characteristics of SEC filings (words per sentence, complex words ratio, file size), the time from announcement to closing (time to close) and the time from the filing date to the closing date (time to close minus time to file), product market competition (using the Hoberg and Phillips (2010) similarity score as computed for the ten nearest neighbors in the product market space), governance characteristics of the filing firm (duality dummy identifying firms in which the CEO is also chairman of the board and the CEO's pay slice from Bebchuk et al. (2011) as measures of CEO power) and its CEO personality trait (CEO overconfidence using the longholder proxy from Malmendier and Tate (2005)). Variables borrowed from prior literature display comparable descriptive statistics in our sample: Hoberg and Phillips (2010) report an average similarity score of 0.201 for the ten nearest neighbors in the product market space, Bebchuk et al. (2011) an average pay slice of 0.357 and using a comparable sample, Aktas et al. (2019) classify twenty-seven percent of CEOs as overconfident using the longholder proxy.

3.4. Econometric specification

We start our multivariate investigations by regressing writing complexity (the Fog index) on investor reactions at the deal announcements (acquirer CAR) and control variables:

$$Fog_i = \alpha + \boldsymbol{\beta_t} + \gamma\ CAR_i^{Acq} + \boldsymbol{Controls_i'}\,\boldsymbol{\delta} + \boldsymbol{IND_i} + \ \epsilon_i \tag{5}$$

where the subscript i is for M&A transaction, $\boldsymbol{\beta_t}$ are year fixed effects, CAR_i^{Acq} are the acquirer CAR, $\boldsymbol{Controls_i}$ and $\boldsymbol{IND_i}$ are respectively the list of control variables and industry dummies. Bold typeface indicates vector notations. Year fixed effects are included to control for changing macro-economic conditions and to account for the positive trend in SEC filing writing complexity documented in Table 1. We estimate Equation 5 by ordinary least square in the main analysis, but assess the robustness of our findings with left censored Tobit specification and an instrumental variable approach. Standard errors are robust to heteroskedasticity. We checked whether our results are potentially affected by multicollinearity issues using the variance inflation factor, but this is not the case (unreported). In robustness checks and additional analyses, we use variations around the baseline specification that will be introduced in the course of the results' presentation. We note also that Fog index limitations when applied to accounting and financial documents pointed out in Loughran and McDonald (2014) do not affect our results because any systematic bias in the left-hand side variable of Equation 5 will be captured by the constant term.

We argue in Section 2 that SEC forms writing is subject to a strict body of rules. If firms strictly comply with these ones, DEFM14A and S-4 filings' writing complexity should be uncorrelated with M&A transactions features and, in particular, with investor reactions at the transaction announcement. Testing whether this is the case or not comes down to test the null hypothesis that the coefficient γ in Equation 5 is equal to zero, which can be performed with a simple test of Student.

4. Results

4.1. Investor reactions feedback on SEC filings writing complexity

We report the estimation results of Equation 5 in Table 4. In each model, the dependent variable is the Fog index (Equation 1) and the independent variable of interest is the acquirer CAR (Equation 3). Column 1 displays univariate results. From columns 2 to 4, we add successively deal level control variables, acquirer and target control variables and the joint filing dummy variable. We report these specifications to check the robustness of our results and in particular, to verify that our results are not affected by the bad control issue (Angrist and Pischke, 2009). The addition of acquirer and target control variables significantly affects our sample size as we lose 398 observations. Checking for the robustness to the joint filing dummy variable appears important because the target decision to also file a separate SEC form potentially signals the presence of some specific information in these filings.

In each specification, the coefficient of acquirer CAR is negative and statistically significant at least at the 5 percent level of confidence: the null hypothesis of absence of correlation between investor reactions at the deal announcement and writing complexity is rejected, suggesting that information collected in SEC filings is not free from endogeneity concern. The coefficient sign indicates that the more negative the investor reactions, the more complex the SEC filings writing. Apparently, acquirers tend to generate more complex filings when investors do not welcome warmly the announced transaction.

We note that coefficient point estimates are stable between columns 1 and 2, an indication that our results are not affected by bad controls. In columns 3 and 4, coefficient point estimates decrease by more than twenty percent (from around -1.9 in columns 1 and 2 to around -2.38 in columns 3 and 4). The loss of one fourth of the sample due to the inclusion of acquirer and target control variables makes however unclear whether this change in coefficient point estimates is driven by a change in sample composition or by the inclusion of additional control variables.

Deal characteristics seem not to affect the complexity of SEC filings writing but bidder and target ones do: the larger the merging parties, the more complex the writing. Target leverage appears also to be positively correlated with writing complexity. We however refrain ourselves to provide any causal interpretation of these coefficients because they are likely to be correlated with latent factors. For example, transactions between larger merging parties may involve more numerous intermediaries (investment banks, legal advisors, accountants, etc.), leading to more complex documents. The inclusion of these control variables helps nevertheless to fight endogeneity due to omitted variable biases when interpreting the coefficient of the acquirer CAR, our independent variable of interest (see the conditional mean independence theorem in Stock and Watson (2020)).

To assess the robustness of the negative relation between acquirer CAR and the Fog index, we perform various checks in Table 5. Panel A replicates the specification in column 4 of Table 4 using alternative dependent variables and estimation methods. We first examine which one of the two components of the Fog index, the number of words per sentence and the complex word ratio, drives the documented negative relation between acquirer CAR and the Fog index. The dependent variable is words per sentence in column 1 and complex word ratio in column 2. The negative relation between acquirer CAR and the Fog index is mainly driven by the number of words per sentence. The Fog index, being censored to the left by construction, we assess the robustness of our findings by adopting a left censored Tobit specification in columns 3 to 5. The Tobit models yield the same conclusion (a negative relation between acquirer CAR and the Fog index, mainly driven by the number of words per sentence). In addition to these robustness checks, we report in Appendix B results obtained replacing the Fog index by seven alternative measures of writing complexity as in Ganguly et al. (2021), namely the Coleman-Liau index, the Flesch Ease Index, the Flesch-Kincaid Readability Index, the RIX Readability Index, the Automated Readability Index, the Smog Readability Index and the Lasbarhets Readility Index. At the exception of the Coleman-Liau index, these are highly correlated to the Fog index and, unsurprisingly, obtained results are highly comparable.

In Panel B of Table 5, we augment our baseline specification with an interaction term between acquirer CAR and a post dummy variable[9] to examine the effect of the Plain Writing Act of 2010 on the sensitivity of the Fog index and the number of words per sentence to investor reactions. The coefficient estimate of acquirer CAR, which captures its effect on the Fog index over the sample period prior to 2011, is negative and statistically highly significant in the reported four models. The interaction term loads with a statistically insignificant positive coefficient, indicating that we cannot reject the null hypothesis that the relation is the same before and after the Plain Writing Act.

4.2. Controlling for file size

Loughran and McDonald (2014) argues that the Fog index is not well suited to evaluate the writing complexity of accounting and financial reports for at least two reasons: words that appear to be complex in the common language are not necessarily so for readers educated in these specialized fields and accounting and financial reports contain many bulleted lists and tables that make difficult (or even irrelevant) the identification of sentences. The authors suggest the use of the file size as an alternative. In Table 3, we report that the file size, computed as the natural logarithm of the HTML-cleaned file size measured in megabyte, is positively correlated with the Fog index and that this correlation is statistically significant at the 1% level. The file size appears therefore, at least in our sample of DEFM14A and S-4 SEC forms, to be related to the Fog index. But the correlation point estimated is 0.07, suggesting that the file size captures also other dimensions of these documents.

To obtain a better picture of the determinants of the file size, we report in Panel A of Table 6 three regressions of the file size on the acquirer CAR in column 1, the absolute value of the acquirer CAR in column 2 and the target firm size in column 3, respectively. We chose these three specifications because they depict respectively how the transaction is welcomed by the acquirer shareholders, the significance of the information release at the announcement date and a mixture between economic relevance and complexity of the acquisition. These regressions include also year and industry fixed effects as in Equation 5. The file size appears to be unrelated to acquirer CAR, but positively and statistically significantly related to the absolute value of acquirer CAR and to the target size: when

[9] Post is a dummy variable identifying the period post to the Plain Writing Act (i.e., year 2011 and onward).

information released at the announcement generates more significant investor anticipations revision and when acquiring bigger targets, DEFFM14A and S-4 forms filed to the SEC are longer. The file size, in addition to the writing complexity, is capturing these additional deal characteristics.

It is therefore interesting to test whether our baseline results are robust to the inclusion of the file size as an additional control variable. To this end, we augment the specifications in Table 4 with the file size as an additional control variable. Results are reported in Panel B of Table 6. We start by noting that the coefficient of file size is negative and highly significant in all specifications. Moreover, the coefficient point estimates are stable across all four specifications. This comes out as a surprise because the bivariate evidence (the correlation coefficient reported in Table 3) indicates a positive relation between the Fog index and the file size. Apparently, controlling for acquirer CAR, year dummies and industry fixed effects partials out drivers of this positive correlation and, conditionally on these control variables, longer filings better conform with simpler writing. Next, we observe that the results in Table 4 are unaffected by the inclusion of the file size as an additional control variable. Coefficient point estimates are comparable and, if anything, statistical significances are higher. The feedback effect of investor reactions at the transaction announcement on the writing complexity is not driven by the additional transaction characteristics proxied by the file size, a reassuring finding.

4.3. Investor reactions feedback on SEC filings time to file

We argue in Section 3 that acquirers have significant leeway in their choice of the time to file. It is therefore interesting to investigate whether this decision is also affected by investor reactions at the deal announcement. The presence of such additional feedback effect would reinforce violation suspicions of SEC filings' content exogeneity assumption. We undertake this investigation by replicating Equation 5 regression specification but with the logarithm of the time to file as left-hand side variable. We apply the logarithmic transformation to account for the count data nature of the time to file variable.

Table 7 provides the results, adopting the same presentation structure as in Table 4. The coefficient estimate of acquirer CAR is positive and highly statistically significant in all four specifications: the more investors react positively to the transaction announcement, the longer the time to file. Transactions warmly welcomed by investors are not only associated with simpler SEC forms, but they are also filed later. To the extent that the time to file captures an additional discretionary acquirer choice, the evidence supports again that acquirers account for investor reactions when writing and filing SEC forms.

One may however question whether the time to file is a significant issue in itself. Isn't it the time from announcement to closing that matters for acquirers at the end? To shed some light on this question, we decompose the time to close (i.e., number of days between the announcement and the closing of the corresponding transaction) into the time to file and the difference between the time to close and the time to file. Table 8 replicates the specification in column 4 of Table 7 (the model with the full set of control variables) with these alternative left hand side variables in logarithmic transform. The dependent variable is the time to close in column 1 and the time to close minus time to file in column 2.

Results reported in Table 8 indicate that there is also a feedback effect on the time to close with the coefficient estimate of acquirer CAR being positive and marginally significant. But this effect is clearly driven by the feedback effect on the time to file documented in Table 7, because investor reactions do not affect the difference between the time to close and the time to file (see column 2 of Table 8). Collectively, these results emphasize that the time to file is the key variable impacted by investor reactions feedback, consistently with the decision to file being under direct control of the acquirer. This is the channel through which investor reactions affect the time to close. This result brings support, albeit indirect, to the presence of strategic information disclosure behaviors by acquirers. This represents a stronger form of exogeneity violation than the presence of latent factors because it is directly related to the agent optimization process.

4.4. Controlling for endogeneity

Our multivariate analyses results are exposed to the classic sources of endogeneity, in particular the omitted variable bias and the endogenous measurement error.[10] Even if this later one would play against rejecting the null hypothesis of absence of relation between acquirer CAR, writing complexity and time to file (the attenuation bias), we adopt classic instrumental variables (IV) identification strategy to test whether a causal interpretation of our results is warranted, noting yet that identification is not sufficient in itself to imply causality (Kahn and Withed, 2018).

Our instruments of acquirer CAR, the potentially endogenous variable, are the Amihud illiquidity ratio, lagged by one year, and the acquirer industry average CAR, computed at the 2-digit SIC industry after exclusion of the transaction under focus. Valid instruments must meet essentially two requirements (Angrist and Pischke, 2009): relevance and exclusion. Thanks to the use of two instruments for one endogenous variable, both are testable using, respectively, a Fisher test of joint significance of instruments in a regression of acquirer CAR on instruments and control variables (the first stage regression when using the two-stage least squares (2SLS) estimator) and a Sargan test of overidentification. We report both in support of our identification strategy. Over and above these

[10] Simultaneity and reverse causality appear less likely in the present case. Investor reactions are fundamentally based on anticipations of cash-flow consequences of acquisitions. It is therefore hard to figure out a channel through which SEC filing writing complexity and/or time to file would affect acquirer CAR.

statistical tests, the intuition behind the use of the Amihud illiquidity ratio as an instrument relates to the specificity of our sample. Our focus on acquirer filings constrains the M&A sample to transactions involving stock payment, as filing to the SEC is not compulsory for the acquirers in full cash transactions. This has a positive side effect, as it allows us to consider the illiquidity ratio as a potential instrument in this very specific case. Arbitrageurs are quite active on shorting acquirer shares when the payment method involves stock, and the more liquid are the acquirer shares, the more they will be active (liquidity facilitates arbitrage, see, e.g., Roll et al., 2007). Illiquidity of the equity share is therefore a firm characteristic correlated with acquirer CAR, through the arbitrage channel (the relevance condition). Moreover, we see no mechanism generating a direct relation between illiquidity and SEC filling writing complexity (the exclusion restriction), especially as we lag by one year the acquirer Amihud illiquidity ratio. Similarly, it is also difficult to conceive a direct link between the average acquirer CAR in the industry of the acquirer and the SEC filing complexity of the focal transaction.

The last panel in Table 2 provides descriptive statistic. The average Amihud illiquidity ratio is 0.459 for our sample of acquiring firms, higher than in Amihud (2002) who report an average of 0.337. The latter sample is however composed only of NYSE firms while close to half of our sample of acquirers is listed on the NASDAQ, and NASDAQ firms are known to display higher Amihud illiquidity ratios (Brennan et al., 2013). The acquirer industry CAR average is -2.4%, a figure in the order of magnitude of previously reported results for acquisitions of public targets paid in stocks (Fuller et al., 2002).

We obtain IV based results using the 2SLS estimator. Results are reported in Table 9. Column 1 reports the first-stage estimates. Both the Amihud illiquidity ratio and the Industry CAR coefficients are highly significant, leading to a Fisher test of joint significance of 10.36, crossing the threshold usually admitted for relevance. In column 2 and 3, we report the second-stage estimates corresponding respectively to column 4 of Tables 4 and 7. Our 2SLS estimates confirm our previous results, alleviating concerns of significant bias due to endogeneity. We note also that the Sargan test of overidentification fails to reject the null hypothesis, supporting the validity of the exclusion restriction.

Comparison of point estimates (-13.57 versus -2.38 for writing complexity and 2.92 versus 0.61 for time to file) may however raise concerns, as discussed in Jiang (2017). The 2SLS estimator is indeed the ratio of the second stage to the first stage and in case of tiny first stage estimates (weak instruments), this leads potentially to significant overestimated causal effects. We note however that, in the present case, the Fisher joint test of IV significance strongly reject the weak instruments null hypothesis and that the coefficient estimates of our IV are not close to zero. The large coefficient estimates of the instrumented acquirer CAR in columns 2 and 3 of Table 9 appear therefore to reflect more that these estimates are relevant for the sub-sample of compliers.[11]

4.5. Additional tests

This section reports three additional tests. We first examine potential factors moderating the intensity of the documented feedback effect. Then, we analyze whether the writing complexity and the time to file conveys valuable information to investors around the filing date. Finally, we repeat our main analysis with the target side filings as a further robustness check.

Is there firm or CEO specific characteristics that influence the sensitivity of the Fog index and the time to file to investor reactions uncovered in Tables 4 and 7? We provide a first exploration focusing on the intensity of product market competition to which the firm is exposed (as an external governance control mechanism), the separation between CEO and chairman functions (as an internal governance control mechanism), the CEO power within the organization and the CEO overconfidence (to capture some CEO behavioral trait). Our respective proxies are the Hoberg and Phillips (2010) similarity score, the classic CEO duality dummy variable, the CEO pay slice introduced in Bebchuck et al. (2011) and the longholder dummy variable as in Malmendier and Tate (2005). These proxies are introduced and described in Section 3.3.

For each of these proxies, we augment the regression specification introduced in Equation (5) with the considered proxy and its interaction term with acquirer CAR. Results are reported in Panel A of Table 10 for the writing complexity and in Panel B for the time to file. At the exception of one (column 2 in Panel A of Table 10), none of the coefficient estimate of the interaction term is statistically significant. The exception (CEO duality for writing complexity) is moreover weakly significant (p-value = 0.09) and would certainly not resist to a Bonferroni (1936) or Romano and Wolf (2005) adjustment for multiple hypotheses test. Apparently, the investor reaction feedback effects on writing complexity and time to file is pervasive and appears not to be mitigated by strong governance and/or CEO characteristics.

The next additional test is whether the writing complexity and timing of the SEC filings convey valuable information to investors around the filing date. To this end, we start by computing the 3-day acquirer abnormal return around the filing date, adopting the same event study method as the one around the announcement date. The average acquirer CAR around the SEC filing date is indistinguishable from zero, as reported in Table 2, suggesting that filings on average do not reveal new information to investors at that time. However, the first and third quartile of the sample

[11] In the LATE theorem vernacular (Imbens and Angrist, 1994), compliers are observations that react to the instrument at the first stage.

distribution reported in Table 2 suggests on their side the presence of a high cross-sectional heterogeneity that could potentially indicate the presence of a value effect associated with information disclosure. We regress therefore the acquirer CAR around the filing date on the Fog index and the time to file, with in addition the same set of control variables and fixed effects as in Equation 5.

Results are reported in Table 11. In columns 1 and 2, the dependent variable is the 3-day acquirer CAR centered on the filing date itself. Next, we replace the acquirer CAR by a dummy variable equal to one if the acquirer CAR is in the first quartile of the sample distribution (columns 3 and 4) or in its fourth quartile (columns 5 and 6). These alternative specifications are designed to test whether the writing complexity and/or the time to file have value effects more specifically on the two extreme quantiles of the CAR distribution. In no specification are the Fog index nor the time to file variable coefficients statistically significantly different from zero: no value effect is associated with the writing complexity and/or the timing of the filing to the SEC. This absence of statistically significant results, assuming they are not the byproduct of a test missing statistical power, suggests that either investors anticipate information disclosure behaviors before the filing date or they appear not to be sensitive to these behaviors.

Finally, we consider target side filings for robustness purposes. As explained in Section 3.2, we replicate the procedure used to collect SEC forms filed by acquirers to the target side. Doing so, we obtain 849 DEFM14A and 22 S-4 forms, for a total of 871 M&A transactions. Out of these, we are able to collect the needed information for 553 observations. Using the sample of target filings, we reproduce the specifications in column 4 of Tables 4 (writing complexity) and Table 7 (time to file) using the target CAR at the announcement day (obtained on a centered three-day event window) as independent variable of interest. Results are summarized in Table 12, column 1 for the Fog index and column 2 for the time to file. Column 1 confirms the presence of a feedback effect from investor reactions to the writing complexity: the lower the target CAR, the more complex the SEC filing writing, like for acquirer CAR (Table 4). However, we do not find a statistically significant relation between target CAR and the time to file. Our sample size being divided by more than two, this absence of statistical significance is potentially due to a too significant loss of statistical power.

The target side filings allow us also to examine whether there is a relation between the bid premium (i.e., the offer price per share, scaled by the target stock price four weeks prior to the announcement day) and the writing complexity and timing of the corresponding SEC filing. Reporting results obtained with the four-week bid premium appears insightful because, in comparison to announcement returns, bid premiums are less subject to rumors, anticipations and/or private information disclosure unrelated to the M&A transaction itself (Eckbo, 2009). Columns 3 and 4 report the results on the sensitivity of the writing complexity and the time to file to the bid premium, respectively. The bid premium results parallel the ones on target CAR. In particular, the writing complexity decreases with the bid premium. When target shareholders receive lower bid premium, the associated SEC filing appears to be less readable.

5. Conclusion

SEC filings have become a major data source in empirical corporate finance thanks to the combination of easy public access via EDGAR, the development of textual analysis algorithms and the ever growing computing power of modern information processing systems. Our understanding of firm behaviors and their value effects has greatly beneficiated from this trend. M&As, product market competition, corporate culture, CEO personality traits, financial constraints, innovation, exposure to cyber-attacks are among the many topics explored in recent contributions with information extracted from SEC filings.

But firms are optimizing economic agents, a main source of endogeneity in social science, and information is valuable. We question therefore whether information collected into SEC filings complies with exogeneity requirements necessary for causal interpretation of regression results.

Private information is, by definition, not observable and therefore, our empirical strategy relies on relating observable SEC filings characteristics to investor reactions to M&A announcements. This setup presents attractive features because M&A transactions are observable and well-identified events for which announcement dates are readily available, the event study method is the well accepted procedure to capture investor reactions to such announcements, M&A SEC filings are well defined and there is a significant time delay between the announcement date and the filing date, allowing to test for the presence of a feedback effect. We focus essentially on the SEC form writing complexity, as measure by the Fog index introduced in Gunning (1952), because SEC form writing is strictly regulated under the Plain Writing Act of 2010 (noting moreover that as soon as in 1998, the SEC introduced a handbook providing guidelines on form writing). As additional evidence, we also investigate the time it takes for the acquirer to file SEC filings into EDGAR.

Our results are unambiguous and strongly reject the null hypothesis of absence of relation between investor reactions to M&A announcement, SEC form writing complexity and the time to file. These results are robust to the inclusion of many control variables, year dummies and industry fixed effects and endogeneity. Writing complexity is negative function of investor reactions and the time to file a positive one.

Having established that information disclosed in SEC filing does not respect exogeneity, the next question is to which extent these firm information disclosure behaviors affect results using information collected in SEC filings. We believe that there isn’t any general and univocal response to this and that potential biases originating from this source of endogeneity are context specific. It appears to us however important to be aware that information extracted from SEC filings are not just like facts and that searching for identification strategies may reveal necessary to uncover causal relations.

Appendix A. Variable definitions.

Unless explicitly mentioned otherwise, Thomson Reuters SDC is the data source for transaction related variables, CRSP for stock market related variables, Compustat for accounting variables, ExecuComp for CEO related variables, and EDGAR for SEC filings related variables.

Dependent variables

Fog index: Equal to 0.4*(average number of words per sentence + percent of complex words). High values of the Fog Index imply less readable text.

Time to file: The number of days from the announcement date of the merger agreement until the filing date into EDGAR. Regression analyses use the logarithm of one plus the variable.

Time to close: The number of days from the announcement date of the merger agreement until its closing. The Regression analyses use the logarithm of one plus the variable.

Time to close minus time to file: The number of days from the filling date into EDGAR until the closing date of the merger. Regression analyses use the logarithm of one plus the variable.

File size: the logarithm of the file size in megabyte. The used filing is the HTML tag cleaned file. Deleting HTML tags is the only modification to the filing.

Words per sentence: The number of words per sentence is computed as the total word count in the cleaned filing over the total number of sentences. To clean the filing, we delete all numbers, spell common abbreviations in full and delete all HTML tags. The sentences are identified by searching for dots, question marks and exclamation marks. Regression analyses use the logarithm of one plus the variable.

Complex word ratio: Complex words are defined as words consisting of three or more syllables. To identify complex words, we use the 2020 Master Dictionary of Loughran and McDonald and filter the word list according to the syllables. The complex word ratio is the number of complex words in the filing over the total number of words in the filing. (Source: EDGAR, Loughran and McDonald Master Dictionary[12])

Acquirer CAR around filing: Cumulative abnormal return for the acquiring firm over the 3-day event window around the filing date to the SEC. The abnormal return is computed by subtracting the return of the equal-weighted CRSP market index from the return of the corresponding firm.

Q1 dummy CAR around filing: Dummy variable equals 1 if the Acquirer CAR is in the lowest quartile of the acquirer CAR distribution, and 0 otherwise.

Q4 dummy CAR around filing: Dummy variable equals 1 if the Acquirer CAR is in the highest quartile of the acquirer CAR distribution, and 0 otherwise.

Independent variables of interest

Acquirer (target) CAR: Cumulative abnormal return for the acquiring (target) firm over the 3-day event window (–1, +1) around the announcement day. The abnormal return is computed by subtracting the return of the equal-weighted CRSP market index from the return of the corresponding firm.

Acquirer CAR below median: Interaction variable between acquirer CAR and a dummy identifying acquirer CAR below the median value of the acquirer CAR distribution.

Acquirer CAR above median: Interaction variable between acquirer CAR and a dummy identifying acquirer CAR below the median value of the acquirer CAR distribution.

Abs(acquirer CAR): Absolute value of acquirer CAR.

[12] https://sraf.nd.edu/textual-analysis/resources/#Master%20Dictionary

Bid premium: The four-week bid premium in percent reported in SDC.

Control variables

Filings per deal: Number of S-4 and DEFM14A filings per deal filed by the acquirer and the target.

Tender offer dummy: Binary variable that equals 1 if the transaction is classified as a tender offer in SDC, 0 otherwise.

Unsolicited dummy: Binary variable that equals 1 if the transaction is classified as unsolicited in SDC, 0 otherwise.

All cash dummy: Binary variable that equals 1 if the payment method is fully in cash.

Toehold: Acquirer's ownership in the target prior to the merger announcement.

Relative deal size: Deal value reported by SDC scaled by acquirer's market value of equity four days prior to the announcement.

Firm size: Natural logarithm of total assets.

Tobin's Q: Total assets plus market value of equity minus book value of equity divided by total assets.

Leverage ratio: Long- and short-term debt divided by total assets.

Joint filings: Binary variable that equals 1 if a separate S-4 or DEFM14A filing is also filed by the other party to the transaction, 0 otherwise.

Similarity score: The average of the ten most similar competitors of a firm using the Text-based Network Industry Classifications (TNIC). (Source: Hoberg and Phillips, 2010)

Duality: Binary variable that takes the value of 1 if the CEO is also the chairman of the board.

Pay slice: The fraction of the aggregate compensation of the top-five executive team captured by the CEO.

Longholder: Binary variable identifying overconfident CEOs. It takes value of 1 during the tenure of a CEO when at least once during the period 2006-2019, the CEO holds an option until the year of expiration, even though the stock option is at least 40% in-the-money entering its final year, and 0 otherwise. CEOs that never exercise options do not reveal beliefs and, thus, longholder is set to missing.

Instruments

Amihud illiquidity ratio: The yearly average of the firm's daily ratio of absolute return to dollar volume, multiplied by 10^6 for proper display as in Amihud (2002). This variable is lagged by one year relative to the transaction announcement date, and multivariate regression analyses use the logarithm of one plus the variable.

Industry CAR: The average bidder CAR in the 2-digit SIC industry of the transaction under focus. The focal acquirer's CAR is excluded from the computation of the industry CAR.

Appendix B. Additional Alternative Readability Measures

Following Ganguly et al. (2021), we consider the following seven alternative readability measures:

(i) *Coleman-Liau Index*: Equal to 5.88*(total number of characters divided by the total number of words) – 29.6*(total number of sentences divided by total number of words). High values imply less readable text.

(ii) *Flesch Reading Ease Index*: Equal to 206.835 – 1.015* (average number of words per sentence) – 84.6* (average number of syllables per word). High values imply easier to read text.

(iii) *Flesch-Kincaid Readability Index*: Equal to 0.39*(average number of words per sentence) + 11.8*(average number of syllables per word) - 15.59. High values of the index imply less readable text.

(iv) *RIX Readability Index:* Equal to number of words with more than six letters divided by the total number of sentences. High values imply less readable text.

(v) *Automated Readability Index*: Equal to 4.71*(average number of characters per word) + 0.5 (average number of words per sentence) - 21.43. High values imply less readable text.

(vi) *Smog Readability Index:* Equal to 1.043 * Square root of (total number of complex words * 30 divided by total number of sentences) + 3.1291. High values imply less readable text.

(vii) *Lasbarhets Readability Index:* Equal to average number of words per sentence + number of words with more than six letters multiplied by 100 divided by total number of words. High values imply less readable text.

Table B.1 reports summary statistics for these alternative readability measures. . Table B.2 reports the correlation coefficients between the Fog index, the considered alternative readability measures and their corresponding level of statistical significance Table B.3 replicates the specification in column 4 of Table 4 with the alternative readability measures as dependent variable.

Table B.1. Summary Statistics

Table B.1 reports summary statistics for the considered alternative readability measures. N is the number of observations, Mean the arithmetic average, Std. dev. the standard deviation, and Q1, Q2 and Q3, the corresponding three quartiles.

	N (1)	Mean (2)	Std. dev. (3)	Q1 (4)	Q2 (5)	Q3 (6)
Coleman-Liau Index	1,698	31.164	0.714	30.735	31.084	31.480
Flesch Reading Ease Index	1,698	31.463	6.921	26.744	31.319	35.757
Flesch-Kincaid Index	1,698	17.301	2.341	15.795	17.341	18.828
RIX Readability Index	1,698	10.791	1.824	9.648	10.787	11.939
Automated Readability Index	1,698	21.223	2.745	19.489	21.263	23.068
Smog Readability Index	1,698	18.531	1.398	17.678	18.591	19.444
Lasbarhets Readability Index	1,698	65.778	5.736	62.266	65.827	69.439

Table B.2. Correlation Matrix

Table B.2 reports correlation coefficients between the considered readability measures and their corresponding level of statistical significance. ***, **, and * indicate significance at 1%, 5% and 10%, respectively.

	Fog Index	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Coleman-Liau Readability Index (i)	-0.034						
Flesch Reading Ease Index (ii)	-0.949***	-0.065***					
Flesch-Kincaid Index (iii)	0.994***	-0.039	-0.961***				
RIX Readability Index (iv)	0.975***	0.060**	-0.924***	0.971***			
Automated Readability Index (v)	0.976***	0.095***	-0.907***	0.976***	0.973***		
Smog Readability Index (vi)	0.992***	0.017	-0.962***	0.983***	0.967***	0.958***	
Lasbarhets Readability Index (vii)	0.978***	0.056**	-0.926***	0.974***	0.999***	0.977***	0.969***

Table B.3. Investor Reactions and Writing Complexity – Alternative Readability Measures

Table B.3 replicates the Table 4 column 4 specification with alternative readability measures as dependent variable. All columns report the coefficient estimates obtained using the OLS estimator. Acquirer CAR definition is provided in Appendix A. Robust standard errors are reported in parentheses below the coefficient estimates. All models include the same control variables as in column 4 of Table 4, an intercept term, year dummies and industry fixed effects (FE), whose coefficients are not reported for brevity. Year and industry FE are based on calendar year and 2-digit SIC codes classification dummies. ***, **, and * indicate significance at 1%, 5% and 10%, respectively.

	Coleman-Liau Index (1)	Flesch Reading Ease Index (2)	Flesch-Kincaid Index (3)	RIX Readability Index (4)	Automated Readability Index (5)	Smog Readability Index (6)	Lasbarhets Readability Index (7)
Acquirer CAR	-0.451*	8.031***	-2.519***	-2.080***	-3.078***	-1.482***	-6.425***
	(0.252)	(2.385)	(0.829)	(0.648)	(0.987)	(0.505)	(2.001)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,291	1,291	1,291	1,291	1,291	1,291	1,291
R-squared	0.139	0.362	0.378	0.395	0.370	0.391	0.388

References

Ahern, K. R., Sosyura, D., 2014, Who writes the news? Corporate press releases during merger negotiations, Journal of Finance 69, 241-291.

Aktas, N., de Bodt, E. Roll, R., 2010, Negotiations under the threat of an auction, Journal of Financial Economics 98, 241-255.

Aktas, N., de Bodt, E., Bollaert, H., Roll, R., 2016, CEO narcissism and the takeover process: From private initiation to deal completion, Journal of Financial and Quantitative Analysis 51, 113-137.

Aktas, N., Louca, C., Petmezas, D., 2019, CEO overconfidence and the value of corporate cash holdings, Journal of Corporate Finance 54, 85-106.

Alexandridis, G., Antypas, N., Travlos, N., 2017, Value creation from M&As: New evidence, Journal of Corporate Finance 45, 632-650.

Amihud, Y., 2002, Illiquidity and stock returns: cross-section and time-series effects, Journal of Financial Markets 5, 31–56.

Angrist, J.D., Pischke, J.S., 2009, Mostly Harmless Econometrics, An Empiricist's Companion, Princeton University Press, New Jersey.

Bartov, E., 1993, The timing of asset sales and earning manipulation, The Accounting Review 68, 840-855.

Baxamusa, M., Jalal, A., Jha, A., 2018, It pays to partner with a firm that writes annual reports well, Journal of Banking and Finance 92, 13-34.

Bebchuk, L.A., Cremers, K.J.M., Peyer, U.C., 2011, The CEO pay slice, Journal of Financial Economics 102, 199-221.

Bergstresser, D., Desai, M., Rauh, J., 2006, Earnings manipulation, pension assumptions, and managerial investment decisions, Quarterly Journal of Economics 121, 157-195.

Betton, S., Eckbo,B.E., Thorburn, K.S., 2008. Corporate takeovers, in Handbook of Corporate Finance: Empirical Corporate Finance, Eckbo, B.E., ed. 2, 291–429, North Holland, Elsevier.

Bodnaruk, A., Loughran, T., McDonald, B., 2015, Using 10-K text to gauge financial constraints, Journal of Financial and Quantitative Analysis 50, 623-646.

Bonferroni, C.E., 1936, Teoria statistica delle classi e calcolo delle probabilità, Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze 8, 3-62.

Boone, A.L., Mulherin, J.H., 2007, How are firms sold?, Journal of Finance 62, 847-875.

Boone, A.L., Mulherin, J.H., 2008, Do auctions induce a winner's curse?, Journal of Financial Economics 89, 1-19.

Brennan, M., Huh, S. W., Subrahmanyam, A., 2013, An analysis of the Amihud illiquidity premium, The Review of Asset Pricing Studies, 3, 133-176.

Brown, S.J., Warner, J.B., 1985, Using daily stock returns: The Case of event studies, Journal of Financial Economics 14, 3-31.

Buehlmaier, M.M.M., Whited, T.M., 2018, Are financial constraints priced? Evidence from textual analysis, Review of Financial Studies 31, 2693-2728.

Caskey, J., Laux, V., 2017, Corporate governance, accounting conservatism, and manipulation, Management Science 63, 424-437.

Crocker, K.J., Slemrod, J., 2007, The economics of earnings manipulation and managerial compensation, RAND Journal of Economics 38, 698-713.

Cunningham, S., 2021, Causal Inference: The Mixtape, Yale University Press, New Haven and London.

Dyer, T., Lang, M., Stice-Lawrence, L., 2017, The evolution of 10-K textual disclosure: Evidence from Latent Dirichlet Allocation, Journal of Accounting and Economics 64, 221-245.

Eaton, G.W., Guo, F., Liu, T., Officer, M.S., 2021, Peer selection and valuation in mergers and acquisitions, Journal of Financial Economics, forthcoming.

Eckbo, B.E., 2009, Bidding strategies and takeover premiums: A review, Journal of Corporate Finance 15, 149-178.

Erickson, M., Hanlon, M., Maydew, E.L., 2006, Is there a link between executive equity incentives and accounting fraud?, Journal of Accounting Research 44, 113-143.

Fama, E.F., Fisher, L., Jensen, M.C., Roll, R., 1969, The adjustment of stock prices to new information, International Economic Review 10, 1-21.

Florackis, C., Louca, C., Michaely, R., Weber, M., 2020, Cybersecurity Risk, University of Chicago, Becker Friedman Institute for Economics Working Paper No. 2020-178, Available at SSRN: https://ssrn.com/abstract=3748622.

Fuller, K., Netter, J.M., Stegemoller, M., 2002, What do returns to acquiring firms tell us? Evidence from firms that make many acquisitions, Journal of Finance 57, 1763–1793.

Ganguly, A., Ganguly, A., Ge, L., Zutter, C. J., 2021, Shareholder litigation risk and readability of corporate financial disclosures: Evidence from natural experiments, Available at SSRN: https://ssrn.com/abstract=3401891

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

Hansen, R.G., 2001, Auctions of Companies, Economic Inquiry 39, 30-43.

Harford, J., 2005, What drives merger waves? Journal of Financial Economics 77, 529–560.

Hoberg, G., Phillips, G., 2010, Product market synergies and competition in mergers and acquisitions: A text-based analysis, Review of Financial Studies 23, 3773–3811.

Hoberg, G., Maksimovic, V., 2015, Redefining financial constraints: A text-based analysis, Review of Financial Studies 28, 1312-1352.

Hoberg, G., Phillips, G., 2016, Text-based network industries and endogenous product differentiation, Journal of Political Economy 124, 1423-1465.

Imperatore, C., Pündrich, G., Verdi, R.S., Yost, B.P., 2021, The strategic choice of peers in M&A valuations, Working paper, MIT Sloan School of Management, Available at: https://ssrn.com/abstract=3869109.

Jiang, W., 2017, Have Instrumental Variables Brought Us Closer to the Truth, Review of Corporate Finance Studies 6,127–140

Kahn, R., Whited, T., 2018, Identification Is Not Causality, and Vice Versa, Review of Corporate Finance Studies 7, 1-21

Kau, J.B., Linck, J.S., Rubin, P.H., 2008, Do managers listen to the market? Journal of Corporate Finance 14, 347-362.

Imbens, G., Angrist, J., 1994, Identification and Estimation of Local Average Treatment Effects, Econometrica 62, 467-476

Li, F., 2008, Annual report readability, current earnings, and earnings persistence, Journal of Accounting and Economics 45, 221-247.

Li, K., Mai, F., Shen, R., Yan, X., 2021, Measuring corporate culture using machine learning, Review of Financial Studies 34, 3265-3315.

Liu, T., Officer, M., 2021, Public news, stock-market reactions, and bidding behavior in private merger negotiations, Working paper, Iowa State University, Available at: https://ssrn.com/abstract=3383209.

Loughran, T., McDonald, B., 2011, When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks, Journal of Finance 66, 35-65.

Loughran, T., McDonald, B., 2014, Measuring readability in financial disclosures, Journal of Finance 69, 1643-167.

Loughran, T., McDonald, B., 2020, Textual Analysis in Finance, Annual Review of Financial Economics 12, 357-375.

Luo, Y., 2005, Do insiders learn from outsiders? Evidence from mergers and acquisitions, Journal of Finance 60, 1951–1982.

Malmendier, U., Tate, G., 2005, CEO overconfidence and corporate investment, Journal of Finance 60, 2661–2700.

Masulis, R.W., Simsir, S.A., 2018, Deal Initiation in mergers and acquisitions, Journal of Financial and Quantitative Analysis 53, 2389-2430.

Mukherjee, A., Singh, M., Zaldokas, A., 2017, Do corporate taxes hinder innovation? Journal of Financial Economics 124, 195-221.

Mulherin, J.H., Simsir, S.A., 2015, Measuring deal premiums in takeovers, Financial Management 44, 1-14.

Officer, M.S., Poulsen, A.B., Stegemoller, M., 2009, Target-firm information asymmetry and acquirer returns, Review of Finance 13, 467-493.

Roll, R., Schwartz, E., and Subrahmanyam, A., 2007, Liquidity and the law of one price: The case of the futures-cash basis, Journal of Finance 62, 2201-2234.

Romano J.P., Wolf M., 2005, Exact and approximate stepdown methods for multiple hypothesis testing, Journal of the American Statistical Association 100, 94-108.

Stock, J.H., Watson, M.W., 2020, Introduction to Econometrics, Fourth Edition, Pearson Global Edition.

Strobl, G., 2013, Earnings manipulation and the cost of capital, Journal of Accounting Research 51, 449-473.

Figure 1 – Number of DEFM14A form SEC filings by year

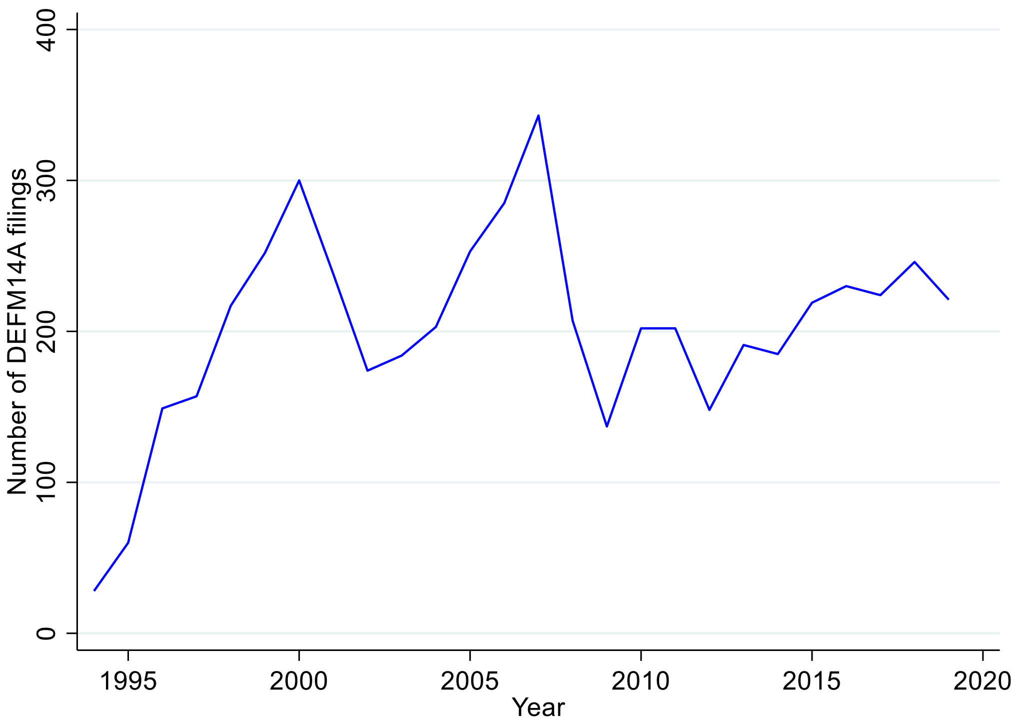

Table 1. Sample Distribution by Announcement Year

This table presents the yearly distribution of our sample, which includes deals announced over the 1994-2019 period and with available S-4 or DEFM14A SEC filings filed by the acquirer. Column 1 reports the number of deals (N). Column 2 presents the aggregate deal value in US$ million. Column 3 reports the average Fog index of the corresponding SEC filings. Columns 4 and 5 display the average time to file (i.e., number of days from the announcement date till the filing date) and time to close (i.e., number of days from the announcement date till the closing date), respectively. Variable definitions are in Appendix A.

Year	N (1)	Deal value (2)	Fog index (3)	Time to file (4)	Time to close (5)
1994	32	20,457.2	21.4	128.6	215.3
1995	79	89,180.5	21.2	89.2	169.0
1996	94	131,526.5	21.5	76.6	159.8
1997	156	157,049.8	22.1	74.4	143.9
1998	155	583,433.3	21.9	84.1	172.6
1999	164	320,769.6	21.5	76.1	157.4
2000	124	304,612.2	21.3	54.3	142.1
2001	108	148,043.4	21.1	40.0	124.1
2002	48	80,757.7	21.3	74.0	171.9
2003	72	94,533.9	21.4	55.7	157.7
2004	71	123,936.1	22.2	55.9	167.0
2005	63	262,631.5	22.4	57.2	155.6
2006	48	78,431.5	22.6	56.3	152.9
2007	47	41,998.8	22.4	60.7	160.0
2008	25	42,383.2	22.2	72.5	166.0
2009	32	200,412.9	23.2	55.0	160.9
2010	29	63,548.5	23.0	53.9	199.9
2011	20	28,476.1	23.0	49.1	164.1
2012	26	29,394.2	23.1	70.8	185.8
2013	31	36,987.0	22.6	68.1	173.3
2014	46	167,183.0	23.3	59.4	166.9
2015	52	294,766.9	23.7	55.6	176.5
2016	47	77,771.2	24.4	77.3	178.1
2017	44	145,330.4	24.2	55.7	172.0
2018	54	215,043.1	24.6	49.9	151.5
2019	31	228,898.8	24.7	47.8	148.9
Total	1,698	3,967,557.3	22.6	65.3	165.1

Table 2. Summary Statistics

This table provides descriptive statistics for our sample of deals announced over the 1994-2019 period and with available S-4 or DEFM14A SEC filings filed by the acquirer. The tables report the number of observations (N), the mean, the standard deviation (St. dev.), and the three quartiles (Q1, Q2 and Q3) of the dependent variables, independent variables of interest, and control variables. Variable definitions are in Appendix A.

	N (1)	Mean (2)	St. dev. (3)	Q1 (4)	Q2 (5)	Q3 (6)
<u>*Dependent variables*</u>						
Fog index	1,698	22.197	2.287	20.670	22.191	23.682
Words per sentence	1,698	33.933	5.791	30.117	33.926	37.825
Complex word ratio	1,698	0.216	0.011	0.207	0.215	0.223
Time to close	1,698	159.949	84.399	106.000	144.000	189.000
Time to file	1,698	66.650	53.827	33.000	52.000	85.000
Time to close minus time to file	1,698	93.300	66.343	52.000	80.000	112.000
Acquirer CAR around filing	1,679	0.000	0.042	-0.021	-0.002	0.017
<u>*Independent variables*</u>						
Acquirer CAR	1,698	-0.023	0.070	-0.058	-0.018	0.012
Target CAR	806	0.211	0.209	0.068	0.165	0.316
Bid premium	871	36.510	33.862	15.660	29.150	49.870
Similarity score	1,576	0.259	0.095	0.185	0.247	0.326
Duality	823	0.662	0.473	0.000	1.000	1.000
Pay slice	856	0.387	0.124	0.318	0.389	0.455
Longholder	443	0.284	0.452	0.000	0.000	1.000
<u>*Control variables*</u>						
Filings per deal	1,698	3.600	4.929	2.000	3.000	4.000
Joint filings	1,698	0.203	0.403	0.000	0.000	0.000
File size	1,698	-0.146	0.431	-0.448	-0.189	0.104
Tender offer dummy	1,698	0.027	0.161	0.000	0.000	0.000
Unsolicited dummy	1,698	0.013	0.113	0.000	0.000	0.000
All cash dummy	1,698	0.035	0.185	0.000	0.000	0.000
Toehold	1,698	0.533	4.096	0.000	0.000	0.000
Relative deal size	1,698	0.524	1.723	0.099	0.272	0.638
Bidder firm size	1,628	7.706	2.087	6.402	7.808	9.112
Bidder Tobin's Q	1,592	1.909	1.909	1.054	1.177	1.887
Bidder leverage ratio	1,628	0.206	0.175	0.072	0.162	0.303
Target firm Size	1,415	6.201	1.966	4.762	6.169	7.540
Target Tobin's Q	1,389	1.776	1.653	1.024	1.145	1.764
Target leverage ratio	1,415	0.207	0.204	0.039	0.152	0.318
<u>*Instruments*</u>						
Amihud Illiquidity ratio	1,635	0.459	4.433	0.001	0.004	0.042
Industry CAR	1,500	-0.024	0.042	-0.038	-0.018	-0.007

Table 3. Correlation Matrix

This table shows the Pearson correlation matrix between the dependent and independent variables of interest. Variable definitions are in Appendix A. ***, **, and * indicate significance at 1%, 5% and 10%, respectively.

	Fog index	Time to file	Acquirer CAR	Abs(acquirer CAR)
Time to file	-0.126***			
Acquirer CAR	-0.044*	0.122***		
Abs(acquirer CAR)	-0.030	-0.118***	-0.416***	
File size	0.074***	0.036	-0.011	0.181***

Table 4. Investor Reactions and Writing Complexity

This table presents the effect of investor reactions (i.e., acquirer CAR) on SEC filings writing complexity for our sample of M&A deals announced over the 1994-2019 period and with available S-4 or DEFM14A SEC filings filed by the acquirer. All columns report the coefficient estimates of OLS regressions, with the dependent variable being the Fog index as a proxy for writing complexity. Variable definitions are in Appendix A. Robust standard errors are reported in parentheses below the coefficient estimates. All models include an intercept term, year dummies and industry fixed effects (FE), whose coefficients are suppressed for brevity. Year and industry FE are based on calendar year and 2-digit SIC codes classification dummies, respectively. ***, **, and * indicate significance at 1%, 5% and 10%, respectively.

	(1)	(2)	(3)	(4)
Acquirer CAR	-1.961**	-1.865**	-2.386***	-2.385***
	(0.773)	(0.780)	(0.824)	(0.825)
Filings per deal		0.029***	-0.003	-0.003
		(0.009)	(0.011)	(0.011)
Tender offer dummy		0.619*	0.484*	0.498*
		(0.326)	(0.288)	(0.293)
Unsolicited dummy		0.364	0.377	0.381
		(0.522)	(0.391)	(0.392)
All cash dummy		-0.414	-0.149	-0.144
		(0.289)	(0.288)	(0.288)
Toehold		-0.011	-0.012	-0.013
		(0.012)	(0.013)	(0.014)
Relative deal size		0.001	-0.019	-0.018
		(0.023)	(0.026)	(0.027)
Bidder firm size			0.227***	0.227***
			(0.042)	(0.042)
Bidder Tobin's Q			-0.010	-0.010
			(0.041)	(0.041)
Bidder leverage ratio			0.484	0.491
			(0.368)	(0.368)
Target firm size			0.230***	0.229***
			(0.046)	(0.046)
Target Tobin's Q			0.081*	0.082*
			(0.049)	(0.049)
Target leverage ratio			1.007***	1.015***
			(0.346)	(0.346)
Joint filings				0.055
				(0.151)
Year FE	Yes	Yes	Yes	Yes
Acquirer industry FE	Yes	Yes	Yes	Yes
Observations	1,689	1,689	1,291	1,291
R-squared	0.244	0.251	0.394	0.394
F-statistic	6.44	3.81	19.27	17.87

Table 5. Investor Reactions and Writing Complexity – Robustness Checks
Panel A replicates the specification in column 4 of Table 4 using alternative dependent variables and estimation methods as robustness checks. The first two columns report the coefficient estimates of OLS regressions using the two components of the Fog index as dependent variables. These are words per sentence and complex word ratio, respectively. Columns 3 to 5 report the coefficient estimates of left censored Tobit specifications. The dependent variable is the Fog index in Column 3, and its two components in columns 4 and 5, respectively. Panel B examines the effect of the Plain Writing Act on the sensitivity of the SEC filing writing complexity to investor reactions. The first (last) two columns report the coefficient estimates of OLS (Tobit) regressions using the Fog index and the number of words per sentence as dependent variables, respectively. Post is a dummy variable identifying the period post to the Plain Writing Act (i.e., year 2011 and onward). The single term "Post" is omitted from the specifications because it is subsumed by the inclusion of the intercept term and year dummies. Variable definitions are in Appendix A. Robust standard errors are reported in parentheses below the coefficient estimates. All models include the same set of control variables as in column 4 of Table 4, an intercept term, year dummies and industry fixed effects (FE), whose coefficients are suppressed for brevity. Year and industry FE are based on calendar year and 2-digit SIC codes classification dummies, respectively. ***, **, and * indicate significance at 1%, 5% and 10%, respectively.

Panel A. Alternative models

	Words per sentence OLS (1)	Complex word ratio OLS (2)	Fog index Tobit (3)	Words per sentence Tobit (4)	Complex word ratio Tobit (5)
Acquirer CAR	-5.868***	-0.319	-2.249***	-5.593***	-0.306
	(2.127)	(0.394)	(0.723)	(1.846)	(0.390)
Control variables	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Acquirer industry FE	Yes	Yes	Yes	Yes	Yes
Observations	1,291	1,291	1,301	1,301	1,301
[Pseudo] R-squared	0.374	0.208	[0.126]	[0.085]	[0.080]
F-statistic	19.240	1.814	-	-	-

Panel B. The Effect of the Plain Writing Act of 2010

	Fog index OLS (1)	Words per sentence OLS (2)	Fog index Tobit (3)	Words per sentence Tobit (4)
Acquirer CAR	-2.746***	-6.566***	-2.564***	-6.197***
	(0.921)	(2.367)	(0.803)	(2.030)
Acquirer CAR * Post	2.113	4.085	1.770	3.402
	(1.889)	(5.030)	(1.744)	(4.663)
Control variables	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Acq. Industry FE	Yes	Yes	Yes	Yes
Observations	1,291	1,291	1,301	1,301
[Pseudo] R-squared	0.395	0.374	0.127	0.085
F-statistic	16.720	18.000	-	-

Table 6. Investor Reactions and Writing Complexity – Controlling for File Size

This table examines whether file size is correlated with our independent variable of interest (Panel A) and the robustness of our main findings to the inclusion of file size as an additional control variable in the Fog index regressions (Panel B). File size corresponds to the logarithm of the size of the SEC filings in megabyte. Panel A presents coefficient estimates of OLS regressions where file size is regressed on acquirer CAR in column 1, on the absolute value of acquirer CAR in column 2, and on target firm size in column 3. Panel B replicates Table 4 by augmenting the specification with file size as an additional control variable. Variable definitions are in Appendix A. Robust standard errors are reported in parentheses below the coefficient estimates. All models include an intercept term, year dummies and industry fixed effects (FE), whose coefficients are suppressed for brevity. Year and industry FE are based on calendar year and 2-digit SIC codes classification dummies, respectively. ***, **, and * indicate significance at 1%, 5% and 10%, respectively.

Panel A: File size as dependent variable

	(1)	(2)	(3)
Acquirer CAR	-0.120		
	(0.145)		
Abs(acquirer CAR)		1.416***	
		(0.192)	
Target firm size			0.013**
			(0.006)
Year FE	Yes	Yes	Yes
Acquirer Industry FE	Yes	Yes	Yes
Observations	1,689	1,689	1,405
R-squared	0.303	0.326	0.308
F-statistic	0.684	54.470	4.637

Panel B: Controlling for file size in the Fog index regressions

	(1)	(2)	(3)	(4)
Acquirer CAR	-2.073***	-1.901**	-2.275***	-2.274***
	(0.742)	(0.749)	(0.794)	(0.795)
File size	-0.935***	-0.966***	-0.929***	-0.932***
	(0.137)	(0.138)	(0.163)	(0.164)
Filings per deal		.031***	-0.002	-0.002
		(0.009)	(0.01)	(0.011)
Joint filings				0.080
				(0.149)
Deal characteristics	No	Yes	Yes	Yes
Bidder characteristics	No	No	Yes	Yes
Target characteristics	No	No	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Acquirer industry FE	Yes	Yes	Yes	Yes
Observations	1,689	1,689	1,291	1,291
R-squared	0.2659	0.2736	0.4106	0.411
F-statistic	26.32	10.23	22.85	21.29

Table 7. Investor Reactions and Time to File

This table presents the effect of investor reactions (i.e., acquirer CAR) on the time to file for our sample of M&A deals announced over the 1994-2019 period and with available S-4 or DEFM14A SEC filings filed by the acquirer. The dependent variable is the log of one plus the time to file in all models, with the time to file corresponding to the number of days between the announced date of the deal and the date of filing to the SEC. Variable definitions are in Appendix A. Robust standard errors are reported in parentheses below the coefficient estimates. All models include an intercept term, year dummies and industry fixed effects (FE), whose coefficients are suppressed for brevity. Year and industry FE are based on calendar year and 2-digit SIC codes classification dummies, respectively. ***, **, and * indicate significance at 1%, 5% and 10%, respectively.

	(1)	(2)	(3)	(4)
Acquirer CAR	0.856***	0.791***	0.614**	0.615**
	(0.241)	(0.238)	(0.279)	(0.277)
Filings per deal		-0.007	-0.009*	-0.010*
		(0.005)	(0.005)	(0.005)
Tender offer dummy		-0.935***	-0.853***	-0.829***
		(0.134)	(0.142)	(0.144)
Unsolicited dummy		0.624***	0.703***	0.709***
		(0.167)	(0.159)	(0.159)
All cash dummy		0.037	0.117	0.124
		(0.102)	(0.112)	(0.112)
Toehold		0.002	0.000	0.000
		(0.006)	(0.007)	(0.007)
Relative deal size		0.011	0.007	0.008
		(0.009)	(0.007)	(0.006)
Bidder firm size			-0.056***	-0.055***
			(0.014)	(0.014)
Bidder Tobin's Q			-0.012	-0.012
			(0.011)	(0.011)
Bidder leverage ratio			-0.223*	-0.210
			(0.133)	(0.133)
Target firm size			0.030*	0.028*
			(0.016)	(0.016)
Target Tobin's Q			-0.001	0.000
			(0.014)	(0.014)
Target leverage ratio			-0.061	-0.047
			(0.131)	(0.131)
Joint filings				0.096*
				(0.056)
Year FE	Yes	Yes	Yes	Yes
Acquirer industry FE	Yes	Yes	Yes	Yes
Observations	1,689	1,689	1,291	1,291
R-squared	0.200	0.245	0.289	0.291
F-statistic	12.67	11.90	7.574	7.084

Table 8. Investor Reactions and Time to Close

This table presents the effect of investor reactions (i.e., acquirer CAR) on the time to close for our sample of M&A deals announced over the 1994-2019 period and with available S-4 or DEFM14A SEC filings filed by the acquirer. The dependent variable is the log of one plus the time to close in column 1, with the time to close corresponding to the number of days between the announcement date of the deal and its closing date. In column 2, the dependent variable is the log transform of the time to close minus the time to file. Variable definitions are in Appendix A. Robust standard errors are reported in parentheses below the coefficient estimates. All models include an intercept term, year dummies and industry fixed effects (FE), whose coefficients are suppressed for brevity. Year and industry FE are based on calendar year and 2-digit SIC codes classification dummies, respectively. ***, **, and * indicate significance at 1%, 5% and 10%, respectively.

	Time to close (1)	Time to close minus time to file (2)
Acquirer CAR	0.049*	-0.015
	(0.029)	(0.043)
Filings per deal	0.001	0.002**
	(0.001)	(0.001)
Tender offer dummy	-0.072***	-0.055**
	(0.016)	(0.025)
Unsolicited dummy	0.057***	0.024
	(0.019)	(0.028)
All cash dummy	0.032***	0.026
	(0.011)	(0.020)
Toehold	0.001	0.001
	(0.001)	(0.001)
Relative deal size	0.000	-0.001
	(0.001)	(0.002)
Bidder firm size	-0.007***	-0.004
	(0.001)	(0.002)
Bidder Tobin's Q	-0.004***	-0.004**
	(0.001)	(0.002)
Bidder leverage ratio	-0.035***	-0.043*
	(0.013)	(0.022)
Target firm size	0.012***	0.017***
	(0.002)	(0.003)
Target Tobin's Q	0.000	-0.003
	(0.001)	(0.002)
Target leverage ratio	-0.013	-0.036
	(0.012)	(0.022)
Joint filings	0.000	-0.017*
	(0.006)	(0.009)
Year FE	Yes	Yes
Acquirer industry FE	Yes	Yes
Observations	1,291	1,291
R-squared	0.333	0.221
F-statistic	9.580	5.777

Table 9. Investor Reactions and Writing Complexity – Controlling for Endogeneity

This table replicates columns 4 of Tables 4 and 7 using an instrumental variable approach. Column 1 displays the first-stage regression, and columns 2 and 3 the second-stage regressions for the Fog index and the time to file as dependent variables, respectively. The considered instruments in the first-stage regressions are *Amihud Illiquidity ratio* and *Industry CAR*. Variable definitions are in Appendix A. Robust standard errors are reported in parentheses below the coefficient estimates. All models include the same set of control variables as in column 4 of Tables 4 and 8. Year and industry FE are based on calendar year and 2-digit SIC codes classification dummies, respectively. IV Fisher test is a Fisher test of instrumental variables joint significance. J-test of overidentification is a Sargan test of instrument variables overidentification. ***, **, and * indicate significance at 1%, 5% and 10%, respectively.

	1st stage	2nd stage	
	Acquirer CAR	Fog index	Time to file
	(1)	(2)	(3)
Amihud illiquidity ratio	0.029***		
	(0.009)		
Industry CAR	-0.174***		
	(0.060)		
Acquirer CAR (instrumented)		-13.579***	2.922*
		(5.189)	(1.747)
Control variables	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Acquirer industry FE	Yes	Yes	Yes
Observations	1,100	1,100	1,100
R-squared	0.171	0.280	0.264
IV Fisher test	10.36***		
J-test of overidentification		0.002	0.467

Table 10. Additional Tests – Role of Competition, Governance, and CEO Characteristics

This table examines whether the sensitivity of the Fog index and the time to file to acquirer CAR is affected by product market competition (similarity score), corporate governance characteristic (duality and pay slice), and CEO overconfidence (longholder). The dependent variable is the Fog index in Panel A and the log of one plus the time to file in Panel B. Variable definitions are in Appendix A. Robust standard errors are reported in parentheses below the coefficient estimates. All models include the same set of control variables as in column 3 of Table 4, an intercept term, year dummies and industry fixed effects (FE), whose coefficients are suppressed for brevity. Year and industry FE are based on calendar year and 2-digit SIC codes classification dummies, respectively. ***, **, and * indicate significance at 1%, 5% and 10%, respectively.

Panel A. Fog index regressions

	(1)	(2)	(3)	(4)
Similarity score	-1.151			
	(0.914)			
Similarity score × Acquirer CAR	-11.504			
	(8.964)			
Duality		0.020		
		(0.174)		
Duality × Acquirer CAR		-4.058*		
		(2.394)		
Pay slice			-0.105	
			(0.72)	
Pay slice × Acquirer CAR			8.047	
			(11.135)	
Longholder				-0.081
				(0.263)
Longholder × Acquirer CAR				-0.845
				(3.484)
Acquirer CAR	-2.784***	-3.462***	-2.942**	-2.533
	(0.862)	(1.339)	(1.339)	(1.650)
Control Variables	Yes	Yes	Yes	Yes
Year FE + Acquirer industry FE	Yes	Yes	Yes	Yes
Observations	1,272	722	717	378
R-squared	0.395	0.327	0.381	0.483
F-statistic	16.57	5.220	5.568	2.777

Panel B: Time to file regressions

	(1)	(2)	(3)	(4)
Similarity score	0.354			
	(0.281)			
Similarity score × Acquirer CAR	8.578**			
	(3.513)			
Duality		-0.084		
		(0.062)		
Duality × Acquirer CAR		-0.602		
		(0.797)		
Pay slice			0.038	
			(0.236)	
Pay slice × Acquirer CAR			1.342	
			(3.325)	
Longholder				-0.069
				(0.092)
Longholder × Acquirer CAR				-1.725
				(1.203)
Acquirer CAR	0.878***	0.927**	1.062***	0.287
	(0.281)	(0.391)	(0.404)	(0.619)
Control variables	Yes	Yes	Yes	Yes
Year FE + Acquirer industry FE	Yes	Yes	Yes	Yes
Observations	1,272	722	717	378
R-squared	0.299	0.328	0.323	0.361
F-statistic	6.860	4.281	4.174	3.077

Table 11. Acquirer CAR around the SEC Filing Date

This table presents the estimates of OLS regressions of acquirer 3-day cumulative abnormal returns (CAR) around the date of the filing to the SEC. The sample includes deals announced over the 1994-2019 period and with available S-4 or DEFM14A SEC filings filed by the acquirer. In columns 3 and 4, the dependent variable is a dummy variable identifying acquirer CAR lower than the first quartile (Q1). In columns 5 and 6, the dependent variable is a dummy variable identifying acquirer CAR higher than the third quartile (Q3). Variable definitions are in Appendix A. Robust standard errors are reported in parentheses below the coefficient estimates. All models include the same set of control variables as in column 3 of Table 4, an intercept term, year dummies and industry fixed effects (FE), whose coefficients are suppressed for brevity. Year and industry FE are based on calendar year and 2-digit SIC codes classification dummies, respectively. ***, **, and * indicate significance at 1%, 5% and 10%, respectively.

	Acquirer CAR around Filing (1)	Acquirer CAR around Filing (2)	Q1 dummy CAR around Filing (3)	Q1 dummy CAR around Filing (4)	Q4 dummy CAR around Filing (5)	Q4 dummy CAR around Filing (6)
Fog Index	0.001		0.000		0.004	
	(0.001)		(0.006)		(0.007)	
Log of time to file		0.002		-0.019		-0.001
		(0.002)		(0.021)		(0.020)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Acquirer Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,276	1,276	1,276	1,276	1,276	1,276
R-squared	0.051	0.051	0.104	0.105	0.074	0.074
F-statistic	0.750	0.776	1.626	7.708	0.409	0.392

Table 12. Target Side Filings

This table presents the effect of target CAR and bid premium on the Fog Index and the time to file for our sample of M&A deals announced over the 1994-2019 period and with available S-4 or DEFM14A SEC filings filed by the target. Columns 1 and 3 report OLS regressions with the Fog index as dependent variable. Columns 2 and 4 reports OLS regressions with the log of one plus the time to file as dependent variable. Variable definitions are in Appendix A. Robust standard errors are reported in parentheses below the coefficient estimates. All models include an intercept term, year dummies and industry fixed effects (FE), whose coefficients are suppressed for brevity. Year and industry FE are based on calendar year and 2-digit SIC codes classification dummies, respectively. ***, **, and * indicate significance at 1%, 5% and 10%, respectively.

	Fog index (1)	Time to file (2)	Fog index (3)	Time to file (4)
Target CAR	-0.880**	-0.052		
	(0.443)	(0.113)		
Bid premium			-0.008***	0.001
			(0.003)	(0.001)
Filings per deal	-0.009	0.003	-0.011	0.003
	(0.021)	(0.006)	(0.021)	(0.006)
Tender offer dummy	-1.209**	0.168	-1.276**	0.183
	(0.604)	(0.141)	(0.619)	(0.142)
Unsolicited dummy	0.245	0.366**	0.359	0.348*
	(0.402)	(0.179)	(0.407)	(0.179)
All cash dummy	0.317	-0.239***	0.243	-0.258***
	(0.270)	(0.062)	(0.260)	(0.063)
Toehold	-0.026**	0.012*	-0.027**	0.012*
	(0.013)	(0.006)	(0.013)	(0.006)
Relative deal size	0.021	0.017	0.115	0.026
	(0.129)	(0.050)	(0.128)	(0.048)
Bidder firm size	0.202***	-0.049***	0.211***	-0.041**
	(0.071)	(0.016)	(0.068)	(0.016)
Bidder Tobin's Q	-0.059	-0.032	-0.019	-0.026
	(0.096)	(0.023)	(0.092)	(0.023)
Bidder leverage ratio	-0.903*	-0.052	-0.654	-0.038
	(0.533)	(0.143)	(0.523)	(0.140)
Target firm size	0.099	0.053**	0.086	0.054**
	(0.077)	(0.022)	(0.073)	(0.021)
Target Tobin's Q	0.094	0.001	0.058	-0.004
	(0.086)	(0.020)	(0.082)	(0.02)
Target leverage ratio	0.427	0.014	0.222	-0.036
	(0.519)	(0.132)	(0.461)	(0.122)
Joint filings	-0.135	0.057	-0.177	0.068
	(0.201)	(0.054)	(0.198)	(0.054)
Year FE	Yes	Yes	Yes	Yes
Acquirer industry FE	Yes	Yes	Yes	Yes
Observations	553	553	563	563
R-squared	0.452	0.476	0.453	0.476
F-statistic	3.594	7.020	4.007	7.654