Annual Report Commentary on the Value Creation Process

Vasiliki Athanasakou[‡]

Mahmoud El-Haj[*]

Paul Rayson[*]

Martin Walker[#]

Steven Young[†]

This draft: May 2019

[†]Corresponding author: Lancaster University Management School, Lancaster University, UK, email: s.young@lancaster.ac.uk, [#]Alliance Manchester Business School, Manchester University, UK, [‡]London School of Economics, UK [*]School of Computing and Communications, Lancaster University, UK We are grateful for comments from an anonymous FARS reviewer, Mark Clatworthy, Edward Lee, Ken Lee, Gilad Livne, Thomas Schleicher, Derrald Stice, Jennifer Wu Tucker, and seminar participants at Bristol, Cardiff, Cattolica Del Sacro Cuore, Exeter, HEC Lausanne, King's College London, Lancaster, Leeds, Leicester, Loughborough, Munich, Portsmouth, St Gallen, Strathclyde, Warwick, the 2016 Annual Meeting of the American Accounting Association, the European Accounting Association Annual Congress 2018, and the 12th Annual LSE/AMBS Conference 2018 (Stockholm). Research support was provided by Paulo Alves, José Costa and Steve Wattam. Financial support was provided by the Economic and Social Research Council (ESRC) (contracts ES/J012394/1, ES/K002155/1 and ES/R003904/1), the Research Board of the Institute of Chartered Accountants in England and Wales, and the International Centre for Research in Accounting. Data relating to construction of our strategy wordlist are available via DOI: 10.17635/lancaster/researchdata/232. The method described in the paper for retrieving and analysing narratives from UK annual reports published as PDF files is available as a software tool copyrighted to Lancaster University and free for academic use at http://ucrel.lancaster.ac.uk/cfie/.A version of the dataset employed in the paper (plus an updated version including reports for 2015 and 2016) is available via the same link.

Annual Report Commentary on the Value Creation Process

Abstract

We measure annual report commentary articulating entities' business model and strategy, and then examine the capital market effects of enhancing such disclosure. Our empirical disclosure proxy is based on n-grams drawn from popular strategy textbooks and the academic strategy literature. Validation tests confirm that our score: (a) correlates with manual classifications of the quality of strategy-focused disclosures produced by domain experts; (b) covaries predictably with firm-level drivers of strategy-focused disclosures identified by prior research; and (c) captures the structural break in reporting associated with the regulatory mandate for a subset of London Stock Exchange firms to explain their strategy and business model. Tests using this exogenous and measurable increase in strategy-focused disclosure show that enhanced commentary on strategy and business model is associated with lower investor uncertainty. We also find support for an increase in the speed at which information is incorporated into stock price following the annual report release.

Annual Report Commentary on the Value Creation Process

1. Introduction

Daimler AG's 2017 annual report contains six pages of commentary articulating its strategic objectives and business model for creating and preserving shareholder value. KPMG (2016) highlight a growing international trend for such commentary as policymakers including the US Financial Accounting Standard Board (FASB) (2001), International Accounting Standards Board (IASB) (2010), UK Financial Reporting Council (FRC) (2010a), and European Commission (EC) (2017) promote disclosure in this area. Reflecting this trend, a recent Securities and Exchange Commission (SEC) concepts release aimed at modernizing Regulation S-K disclosure requirements seeks views on whether to revise Item 101(a)(1) to require registrants to describe their business strategy in the Management Discussion and Analysis (SEC 2016: 60). The value of such discourse is nevertheless unclear (EY 2015). We shed light on this evolving area of annual reporting by evaluating the impact of annual report descriptions of the value creation process on the quality of firms' information environment.

Strategy and business model define the value creation process. Information on business models helps investors understand an entity's resource requirements, priorities for action, and prospects (FASB 2001: 14-15, IASB 2010: 12). Business models articulate the system of inputs, activities and outputs required to generate cash flows and create long term value, and are framed by the entity's strategic objectives and plans to achieve them.[1] Business model and strategy reporting is considered a central element of effective annual report commentary (International

[1] Casadesus-Masanell and Ricart (2010) argue that a business model reflects an entity's realized strategy: business model describes the way a firm operates and how it creates value, while strategy refers to the contingent choice of business model. The two concepts coincide in simple competitive situations but are distinct in the presence of important contingencies on which a well-designed strategy is based. We use the term strategy-related reporting hereinafter as shorthand for both aspects, as well as for the value creation process broadly defined. Given the close alignment between strategy and business model we do not attempt to distinguish empirically between the concepts.

Integrated Reporting Council [IIRC] 2011, Guillaume 2018: 7). Market participants view annual report commentary on strategy, business model inputs (e.g., innovation and assets) and business model outcomes (e.g., financial results and value created) as a distinct disclosure element that provides an important contextual framework for interpreting other aspects of financial reporting (FASB 2001, FRC 2016).[2] For example, 78% (64%) of buy-side (sell-side) analysts surveyed by PwC (2017) agree that annual report descriptions of strategy, opportunities, and value drivers impact their work directly. Understanding how firms compete and differentiate themselves is also a key theme highlighted by analysts surveyed by Brown et al. (2016: 145). Management's reluctance or inability to explain the value creation process is predicted to create doubt over leadership quality and may restrict capital flows (FRC 2016, Gu and Li 2007).

Clear and comprehensive corporate reporting on value creation is critical given that over half the analysts surveyed by PwC (2017) highlight the need for more annual report commentary on this topic. However, evidence on the usefulness of value creation descriptions is limited to voluntary and partial disclosure outside the annual report such as innovation plans (Gu and Li 2007), management guidance on specific strategic initiatives such as store openings and schedules for clinical trials (Lu and Tucker 2010), and management presentations (Baginski et al. 2018). The usefulness of annual report disclosures articulating how an entity creates and maintains shareholder value through its business model and strategy therefore remains an open question despite persistent external pressure for firms to provide such information (IASB 2010, FRC 2010a, IIRC 2013, Lev and Gu 2016, EC 2017, PwC 2017).

[2] While it is challenging to fully disentangle commentary on strategy from general corporate reporting, recent guidelines clearly view a formal description of the process for creating and maintaining value as distinct from, and incremental to, standard financial commentary on performance and financial position (European Financial Reporting Advisory Group 2013, IIRC 2013, IASB 2010, FRC 2010a). Consistent with this view, Lev and Gu (2016) propose a strategic resources report as a way for organisations to explain how they create and preserve competitive advantage.

Proponents of business model and strategy descriptions argue that investors require such information to fully process the performance and value implications of accounting results and corporate investment decisions. The importance of investors' preexisting information set as a basis for conditioning their interpretation of new information is supported by theory (Verrecchia 1980, Callen et al. 2013). The contextualizing role of the annual report is also consistent with claims that these documents represent a primary source of non-financial information for investors (IASB 2010, ACCA and Eurosif 2013: 4, IIRC 2013, EY 2015, Brown et al. 2016).[3] Skeptics on the other hand question the value of annual report disclosures generally (Chen and Li 2015: 183, White 2013) and argue that descriptions of business model and strategy are especially prone to bland, boilerplate statements that offer few meaningful insights (EY 2015, Santema and Van De Rijt 2001). The value of such disclosures is particularly pertinent for financial reporting given concerns about unfettered growth in the volume of annual report disclosure and problems posed by the inclusion of irrelevant content (SEC 2013, IASB 2017, FRC 2009, 2011).

We test whether descriptions of the value creation process improve firms' information environment by exploiting a revision in the UK Corporate Governance Code requiring London Stock Exchange (LSE) Main Market firms (but not their Alternative Investment Market counterparts) to explain business model and strategy in their annual report (FRC 2010a). We measure value creation commentary using a comprehensive list of n-grams relating to firm strategy and business model drawn from popular strategy textbooks and the academic strategy literature. Since many words appearing frequently in business model- and strategy-focused commentary also occur naturally in general business reporting (e.g., "acquisition" and

[3] Black Sun Plc reports that 84% of long-term investors use the annual report to evaluate corporate strategy and 53% use it to monitor management credibility and assess whether the board has delivered on its promises (https://www.blacksunplc.com/en/insights/blogs/annual-reports-are-really-very-important-investors-say-so.html).

"industry"), we apply two disambiguation procedures to ensure strategy-related n-grams reflect value creation content (i.e. to reduce Type I errors in our word frequency score). First, we concentrate on annual report sections more likely to contain disclosures on the value creation process such as summary highlights, letter from the board chair and management commentary. Second, we weight each n-gram by the conditional probability that it is predictive of an objectively classified corpus of value creation commentary constructed from unambiguous disclosures that describe strategy, business model, and key performance indicators. N-grams in our list that characterize the value creation corpus receive a weight closer to one in our score whereas those less descriptive of the corpus receive a weight closer (or equal) to zero. Estimates based on our score indicate an 11-page (79%) average incremental increase in value creation annual report commentary for LSE Main Market firms relative to AIM firms in response to the FRC (2010a) disclosure mandate.

We evaluate construct validity by comparing disclosure rankings generated by our score against double-blind classifications of the quality of value creation commentary produced by two domain experts. Our score yields precision and recall rates (Manning and Schütze 1999) that typically exceed 80% for reports classified manually as containing high and low quality strategy-related commentary. We supplement this analysis with large-sample evidence. Causality tests confirm the ability of our score to detect the post-2010 incremental increase in strategy-related annual report disclosure mandated for LSE Main Market firms. Association tests demonstrate that our score correlates with factors predicted by prior research to explain variation in the level of strategy-focused disclosure such as firm-level operational complexity and competition. Consistent with value creation commentary being a distinct dimension of management commentary, our results are robust to controls for other disclosure features including length, forward-lookingness, and Li et al.'s (2013) text-based measure of competition. We also

demonstrate that the 2010 structural break does not load in placebo tests using proxies for total report disclosure, consistent with our score capturing disclosure changes that are distinct from any general growth in annual report content.

We then assess whether annual report commentary on the value creation process helps investors resolve uncertainty by exploiting the exogenous and demonstrable shock to reporting practices for LSE Main Market firms resulting from the 2010 disclosure mandate. Difference-in-differences (DID) tests reveal a statistically and economically significant incremental decrease in bid-ask spread, analyst forecast dispersion, and absolute analyst forecast errors on earnings for LSE Main Market firms (relative to their AIM counterparts) in the post-2010 period. Robustness tests for bid-ask spread using the synthetic control method yield similar findings (Abadie and Gardeazabal 2003). Conclusions also hold using a proxy for firms experiencing the largest increase in our score after the 2010 mandate, suggesting that our DID findings are not attributable to other unobservable regulatory changes that may have occurred in the post-2010 period, although we acknowledge the difficulty of ruling out this possibility entirely. We also document weaker evidence of a post-2010 reduction in stock price delay following the annual report release date (Callen et al. 2013). Finally and as predicted, cross-section tests reveal larger information gains for firms with low analyst coverage prior to the regulatory change. Collectively, our evidence supports the view that annual report descriptions of business model and strategy enhance firms' information environment.

We make several contributions to the literature. We extend evidence on the usefulness of voluntary strategy-related disclosures (Gu and Li 2007, Lu and Tucker 2010, Bagniski et al. 2018) to a more complete description of the value creation process presented in firms' annual reports. Specifically, we focus on managements' detailed articulation of their overall strategic vision and how value is delivered within the framework defined by that vision. Our analysis

therefore speaks directly to claims in the financial statement analysis literature that clarity on firms' competitive environment and approach to operating in that setting provides the foundations for effective accounting and financial analysis (Palepu et al. 2010, Penman 2013).

We also contribute to debate on the usefulness of annual report disclosures. While practitioners stress the importance of the annual report for financial analysis and investment decision-making, research provides mixed evidence on the usefulness of such disclosures. Our analysis helps reconcile these views by studying an aspect of disclosure that aligns naturally with the accepted view of the annual report as a reference document and source of contextual information. Information on value creation provides the framework for understanding how activities and results map into future value; and before the 2010 disclosure mandate UK firms did not articulate their value proposition routinely and comprehensively through other reporting channels. Our evidence that enhanced annual report commentary leads to demonstrable capital market benefits supports the usefulness of this enduring and ubiquitous feature of corporate reporting. In related work, Barth et al. (2016) show that integrated reporting quality correlates with lower information asymmetry proxied by bid-ask spreads. The value creation process is a central theme in integrated reporting (IIRC 2011). Our findings extend Barth et al. (2016) by demonstrating how specific disclosures relating to business model and strategic vision improve the quality of firms' information environment.

Finally, our study is the first attempt of which we are aware to develop a comprehensive measure of strategy-related disclosures using automated scoring methods. We demonstrate that our score captures information beyond Li et al.'s (2013) proxy for competition. Construct validity tests suggest that our score captures aspects of disclosure quantity and quality. The proxy is transparent and simple to implement, broad in nature, and applicable to other disclosure settings such as conference calls and analyst reports.

2. Literature and disclosure practice

2.1 The contextual role of commentary describing the value creation process

Detailed annual report commentary on business model and strategy offers important contextual information to aid outsiders' understanding of the mapping from operating activities into cash flows and ultimately shareholder value (FASB 2001, IASB 2010, FRC 2016). Specifically, descriptions of the value creation process encompassing strategy and business model provide a framework to help investors understand the implications of decisions and events on the timing and magnitude of cash flows. Lev and Gu (2016: 117) emphasize the critical role that contextual information on strategy plays when analyzing complex business organizations subject to competition and fast changing technologies. Reflecting the rising importance of management descriptions of the value creation process, regulators and market participants are increasingly placing strategy and business model at the center of the annual report.

Explicit consideration of the contextual role of information lies beyond the scope of most mainstream disclosure models where management is endowed with private information that will impact firm value when communicated credibly [see Stocken (2012) for a survey]. Rather than conveying direct news about cash flows and firm value, contextualizing commentary on business model and strategy establishes the broader information framework that determines how firm-specific performance snapshots and events affect value. Greater clarity on business model and strategy can therefore enhance the precision of investors' private information set and consequently their ability to evaluate and process new information (Verrecchia 1980). While not necessarily informative in its own right, greater transparency about business model and strategy therefore has the capacity to improve the overall quality of disclosers' information environment in at least two ways. First, it improves the processing ability of investors who were previously unable to access such information themselves, thereby helping to level the information playing

field for non-professional investors. Second, it reduces search costs for investors who previously sought to generate such information privately.

Arguments in the literature concerning the provision of strategy-related commentary are consistent with such disclosure playing a contextual role. For example, clarity on the value creation process is important for identifying key corporate success and risk factors that in turn can help pinpoint critical accounting policies and performance measures (Palepu et al. 2010). Careful analysis of business model and strategy also aids the task of forecasting earnings and cash flows by helping to identify sustainable performance and growth potential (Palepu et al. 2010, Penman 2013). Consistent with these claims, research suggests that analysts view information on strategy as more important than quarterly results (Higgins and Diffenbach 1985). Information on how strategic objectives and planning inform firms' operating and financial plans can also shed light on management quality (Kohut and Segards 1992, Gu and Li 2007).

Supporting the view that strategy-related news lowers investor uncertainty, Lu and Tucker (2012) find that forward-looking signals in strategic plan disclosures contained in earnings press releases are associated with a decrease in bid-ask spreads and that disclosing firms experience an incremental increase in market depth. In addition, Gu and Li (2007) document a positive market reaction to press release disclosures of innovation strategy. Collectively, these findings suggest that investors value forward-looking plans about specific strategic initiatives. In contrast, the role of detailed annual report descriptions of value creation centered on business model and strategy remains unexplored in the literature. Doubt exists over the usefulness of such broad descriptions. First, proprietary costs may constrain the amount of detail managers are willing to provide in the annual report for fear of revealing valuable private information their firm's competitive advantage (Verrecchia 1983, Li et al. 2013). In addition to hiding information, proprietary costs may lead management to obfuscate (bias) disclosures on business model and

strategy to mislead competitors and deter entrants (Li et al. 2013). Second, demand for simplicity coupled with proprietary cost considerations may encourage bland descriptions that offer few new insights or specific details regarding quantitative objectives, how these objectives are monitored, and the action plans needed to realize the desired strategy (Santema and Van De Rijt 2001). Third, to the extent an entity's business model is an implicit, unarticulated set of ideas in the minds of the individuals who lead the organization, there is a risk that disclosures may simply provide a message that managers agree to tell outsiders when asked to explain their value creation approach rather than the fluid and uncertain set of ideas that actually drives the business (Langfield-Smith 1997: 210). Fourth, calls for more disclosure on business models are complicated by the absence of an agreed definition of the concept in the literature (Sinfield et al 2012: 86, European Financial Reporting Advisory Group 2013)

Collectively, these doubts cast a shadow over the value of annual report descriptions of value creation despite increasing pressure on management to report such information. Evidence regarding the usefulness of this growing reporting trend is therefore overdue given concern about the expanding role of the annual report and the costs associated with inclusion of irrelevant disclosures (SEC 2013, IASB 2017).

2.2 Reporting recommendations and requirements

Professional accounting firms (Deloitte 2009) and business associations including the CFA Institute (2006) and the US Chamber of Commerce (2007) argue that firms should provide details on long-term value drivers, initiatives and strategies as an alternative to (quarterly) earnings guidance to help mitigate short-termism and bolster investor confidence in the management. Although Exchange Act Release No 34-48960 recommends the Management Discussion and Analysis section of the 10-K filing includes a discussion of competitive position,

and guidelines on the content of 10-K annual report filings for US registrants refer to "a statement of plans and objectives of management" as an example of information that may be included in forward-looking statements (Ernst and Young 2011: 2-22), the SEC does not require registrants to provide explicit information on corporate objectives, strategy or business model.

The IASB published a non-binding practice statement, effective December 8, 2010, on management commentary promoting disclosures that contribute to an understanding of firms' objectives, strategies for meeting those objectives, and the resources that must be managed to deliver results (IASB 2010, para. 24, 27 and 28). Both the European Financial Reporting Advisory Group (EFRAG) and the IIRC emphasize the need for clear disclosure on business model in firms' annual reports (EFRAG 2013, IIRC 2013). In response, the EC (2017) issued non-binding guidelines for Directive 2014/95/EU (disclosure of non-financial information) that encourage reporting on business models.

From 2006, best practice in the UK saw some firms voluntarily disclosing information on business model and strategy, often within the business review section of the annual report required by the Companies Act 2006 (Section 417) (PricewaterhouseCoopers 2012). An influential House of Commons Treasury Committee report (2009, para. 247) nevertheless argued that most firms failed to provide clear information and called for more detail on management's approach to creating and preserving value. The proposal was formalized in a landmark revision to the UK Corporate Governance Code (FRC 2010a). Specifically, paragraph C.1.2 of the 2010 version of the UK Corporate Governance Code, effective for 12-month reporting periods ending on or after June 28, 2011, states:

> *"… directors should include in the annual report an explanation of the basis on which the company generates or preserves value over the longer term (the business model) and the strategy for delivering the objectives of the company."*

The provisions of the Code apply to all firms with a listing of equity shares on the LSE Main Market.[4] The Code does not prescribe how or where firms should provide such commentary, although a footnote to paragraph C.1.2 recommends locating disclosures in the same section as the business review section. Paragraph C.1.2 represented the only substantive disclosure change in the 2010 of the UK Code Governance Code (FRC 2010b: 10-12).

3. Measuring value creation disclosure

This section explains our approach to measuring annual report commentary on the value creation process relating to business model and strategy (hereinafter, strategy-focused commentary). As strategy and the business model are key to the value creation process, for simplicity in this section we refer to this commentary as strategy related commentary. Our starting point involves deriving a comprehensive list of business model- and strategy-related keywords and phrases. In the absence of an appropriate pre-existing wordlist, we derive an externally defined preliminary list by pooling all non-duplicate words and n-grams from the indexes of following five leading business strategy textbooks: Porter (1985), Barney and Clark (2007), Rumelt (2011), Magretta (2012) and Grant (2013). For completeness, we also include n-grams reported by Ronda-Pupol and Guerras-Martin (2012, Appendices 1a through 1c) in their study examining the evolution of the strategy concept in the academic literature. The initial pooled list comprises 4,384 words and phrases, which reduces to 3,584 following exclusion of proper nouns (firms, products and authors) and other irrelevant content. This list is then curated

[4] Financial Conduct Authority Listing Rules give management the option to comply with the provisions of the Code or explain in their annual report why they have failed to do so. The requirement to highlight and explain non-compliance follows the established governance reporting model in the UK, which has been shown to effect structural change in governance arrangements due to the reputational damage and increased scrutiny resulting from non-compliance (Dahya et al. 2002). While non-compliance is therefore an option for management, in reality compliance rates approach 100 percent. Guidance on matters to consider when explaining business model is provided by the Accounting Standard Board (ASB) (2006, para. 30 to 32).

manually and independently by two members of the research team to: (a) remove generic words such as "profit" and "asset" that are unlikely to discriminate between strategy-related content and other aspects of management commentary; (b) simplify or generalize n-grams and lemmatize keywords;[5] (c) expand the list to include inflections and plurals, and alternative English or American spellings; and (d) add disambiguation conditions. Disagreements between coders are reviewed and reconciled. The final list of strategy n-grams comprises 709 elements that nest 2,907 (81%) of the 3,584 words and phrases in the initial filtered list. Appendix A presents the final n-gram list together with further details of the procedure used to construct it. We also provide full reconciliation of the initial and final lists in an online appendix (doi: 10.17635/lancaster/researchdata/232).

While our manually curated wordlist is constructed with the goal of disambiguation in mind, it is nevertheless likely that some elements will occur frequently in discussions unrelated to the value creation process. For example, words like "acquire" and "acquisition" occur throughout the annual report in non-strategy-related contexts. Failure to disambiguate context and meaning is a widely acknowledged problem associated with a bag-of-words approach (Loughran and McDonald 2016). We apply the following two refinements to minimize risk of false positives. First, we restrict our focus to the subset of annual report sections where value creation commentary is more likely to occur. Specifically, we exclude all report sections that form part of the audited financial statements on the basis that text in these sections relates

[5] For example, the n-gram "market share" is removed because it is subsumed in the more general "market" n-gram that nests alternative phrasing such as "share of market", share of the market", and "share in the market", as well as synonyms such as "market fraction", "fraction of the market", "market take", etc.

primarily to accounting treatments, as well as corporate governance statements and remuneration reports where commentary is shaped to a large degree by regulatory compliance.[6]

Our second disambiguation procedure weights the frequency of the k^{th} n-gram by the conditional probability that it is predictive of strategy-focused commentary:

$$StratScore_j = \sum_k \left[count_{kj} \times \Pr\left(S' \mid keyword_k\right) \right], \quad (1)$$

where *StratScore* is our weighted measure of strategy-related commentary for the j^{th} report, *count* is the frequency count of the k^{th} n-gram, and $\Pr(S' \mid keyword_k)$ is the n-gram-specific conditional probability that element *k* is associated with commentaries that unambiguously contain value creation content. *StratScore* is an unscaled measure of strategy-related commentary because we seek to capture the absolute amount of disclosure rather than the proportion of total disclosure devoted to business model and strategy. We control for scale effects by including a proxy for annual report length in our empirical models (Kravet et al. 2013), but for completeness we also present results estimated using a scaled version of the measure.

N-gram-specific conditional probabilities in equation (1) are derived using the following corpus-based application of Bayes rule. We first construct a corpus of business model and strategy commentary by exploiting headings listed in the tables of contents to identify all sections from our sample of annual reports that unambiguously contain strategy-related commentary. Specifically, we use a java script to search all tables of contents section headers (*S*) for keywords "strategy", "strategies", "strategic", "business model", "key performance indicator", and "KPI".

[6] The consequence of this filtering approach is that it risks understanding the incidence of strategy-related commentary by increasing the probability of Type II errors. Given the disambiguation problems associated with applying wordlists, the benefit of reducing Type I errors likely outweighs the corresponding cost of increasing Type II errors. To assess the sensitivity of our results to our section filtering approach, we repeated the analysis after retaining governance statements and remuneration reports on the basis that they are the next most likely candidates to contain strategy-related commentary. Inferences from tests using this alternative subset of annual report sections are identical to those reported in the main body of the paper.

Section titles that include at least one of these keywords are classified as strategy sections (S'); and pooling content across all S' yields our corpus of strategy-related commentary. We also create a corpus of general annual report narrative commentary by pooling content from all S sections across the report sample. We then apply Bayes rule to derive the conditional probability that n-gram k is a member of S' :

$$\Pr(S' \mid keyword_k) = \frac{\Pr(keyword_k \mid S') \times \Pr(S')}{\Pr(keyword_k)}, \tag{2}$$

where:

$$\Pr(keyword_k) = \sum_{s \in S} count_{sk} \sum_{s \in S} \sum_{\kappa=1}^{K} count_{s\kappa} \tag{2a}$$

$$\Pr(keyword_k \mid S') = \sum_{s \in S'} count_{sk} \Big/ \sum_{s \in S'} \sum_{\kappa=1}^{K} count_{sk} \tag{2b}$$

$$\Pr(S') = \sum_{s \in S'} \sum_{\kappa=1}^{K} count_{sk} \Big/ \sum_{s \in S} \sum_{\kappa=1}^{K} count_{sk} \tag{2c}$$

$\sum_{s \in S} \sum_{k} count_{sk}$ = Count of 709 n-grams in all sections of the annual report corpus;

$\sum_{s \in S} count_{sk}$ = Count of n-gram k in all sections of the annual report corpus;

$\sum_{s \in S'} \sum_{k} count_{sk}$ = Count of 709 n-grams in all strategic sections of the annual report corpus;

$\sum_{s \in S'} count_{sk}$ = Count of n-gram k in all strategic sections of the annual report corpus;

K = set of 709 n-grams;

k = n-gram k, $k = 1 \ldots 709$;

s = annual report section subscript;

S = pooled set of narrative sections in the annual report corpus;

S' = pooled set of strategic narrative sections in the annual report corpus.

Equation (2a) is the probability that n-gram k occurs in the corpus of annual report narratives; equation (2b) is the probability that n-gram k appears in the corpus of strategy-related commentary; and equation (2c) is the probability of the section being strategic when n-gram k is

present. The conditional probability for k given S' is invariant over time and across firms because S' and S are created by pooling across all reports in the sample.

Equation (2) assigns a weight of zero to any n-gram k that never appears in the corpus of strategy-focused commentary. We also set equation (2) to zero for any k that has a lower proportion of occurrences in the strategy corpus relative to the full annual report narratives corpus. Applying these zero weights reduces the number of elements in K from 709 to 231. (See Appendix A, Panel A, for the list of non-zero weighted n-grams.) The high incidence of zero-weighted elements is consistent with the regular occurrence of strategy keywords in non-strategy-specific commentary. Accordingly, the value of $StratScore_j$ from equation (1) measuring the level of strategy-related content in a given annual report is equal to the weighted sum of the 231 non-zero-weighted n-grams from K. Section 5 reports results of tests evaluating the validity of *StratScore* as proxy for annual report commentary on business model and strategy.

4. Sample and data

Annual reports published as PDF files were collected from Perfect Information for all LSE Main Market and AIM non-financial firms for fiscal year-ends between July 1, 2002 and June 30, 2014. The initial Perfect Information sample comprises 15,954 annual reports, from which we exclude regulatory reports, non-English language reports, stand-alone sustainability reports, and annual reports for financial periods less than 11 months or greater than 15 months.[7] The sample window starts in 2002 because the availability of digital PDF reports was more

[7] There is no UK equivalent to either the 10-K report filed by US registrants or the Securities and Exchange Commission's EDGAR system. Instead, UK firms distribute their annual reports as a glossy brochure-style digital PDF file (Lang and Stice-Lawrence 2015). Contrary to the standardized 10-K annual report template for US registrants, no standardized template exists for the narrative component. Instead, management is permitted to locate and structure disclosures as they see fit. This is consistent with the general absence of prescriptive disclosure templates in UK company law and securities law, a consequence of which is that the specific content and format of annual report narratives varies significantly across firms (FRC 2012: 8).

limited before this date. We end the sample window in 2014 to limit confounding affects from subsequent changes to the Companies Act 2006 (Strategic Report and Directors' Report Regulations 2013) that changed the format but not the mandated content of strategy-related disclosures. Sample years run July 1, t to 30 June, $t + 1$ to align with the implementation date for the strategy reporting mandate.

We use the algorithm in El-Haj et al. (2019) to retrieve annual report structure and text using the report table of contents. The retrieval procedure is unable to process 1,656 PDF reports from the initial sample for the reasons described in El-Haj et al. (2019), the primary one being that the PDF is an image-based file. Processed reports are filtered to exclude cases where the retrieval process likely results in material error.[8] This filtering process excludes 3,855 reports. Remaining reports are matched manually with firm-level identifiers from Datastream, resulting in the loss of nine observations for which Datastream codes are unavailable. Removing duplicate reports due to changes in year-end reduces the sample by a further 76 observations. Finally, missing accounting and market data required for our empirical tests reduces the final sample to 9,127 observations. Panel A of Table 1 summarizes the sample selection process while Panel B presents frequency counts of reports by year and firm. Observations are evenly distributed over time with the exception of 2002, where the low frequency is due to a higher incidence of image-based PDF files. The final sample represents approximately 50% of non-financial LSE firms with financial statement and market data on Datastream in any given sample year. Most firms (86 percent) have at least two years of data.

[8] Retrieval errors are identified using the following criteria: (a) the narratives component of the annual report comprises less than two sections or less than four pages or less than 100 words, or does not contain at least one key section (chairman's statement, performance commentary, governance statement or remuneration report); (b) the financial statements component of the report comprises less than four sections or less than five pages or less than 100 words; (c) the total annual report word count exceeds 150,000 words; and (d) the annual report start page determined by our procedure is greater than page eight.

Descriptive statistics for *StratScore* and associated variables based on the sample of 9,127 reports are presented in Panel A of Table 2. Total word count for the mean (median) annual report narrative component excluding the governance statement and remuneration report is 13,291 (9,317), while the corresponding raw strategy n-gram count is 159 (99). Weighting strategy words by their corresponding conditional probability value yields an average (median) *StratScore* value of 21.3 (13) based on the 231 non-zero-weighted n-grams. The median conditional probability weight for these 231 n-grams is 0.14, highlighting the universal nature of many words and phrases that characterize discussions of business model and strategy in the professional and academic literatures.

To aid economic interpretation we translate *StratScore* into an estimate of total strategy-related commentary by regressing annual report narrative page count on *StratScore*.[9] (See Appendix B.) The estimated *StratScore* coefficient mapping strategy-related words into annual report narrative content is 0.68 for the full sample. Applying this mapping coefficient to the sample average *StratScore* value of 21.3 from Table 2 implies that a typical report issued during the sample period contains 15 pages (0.68 × 21.3) of strategy-relevant content, which equates to 53% of the 28 pages of narratives in an average report. Appendix B presents separate mapping coefficients for Main Market and AIM firms using samples of reports issued before revisions to the UK Corporate Governance Code 2010 were implemented. Combining these coefficients with sample average page counts implies an incremental average increase in strategy-related commentary for Main Market firms over their AIM counterparts of 11 pages following

[9] We favour page count over word count because casual empiricism suggests that firms provide significant strategy-related material in infographics and pictures. These potentially important aspects of reports are overlooked using word count. Insights are qualitatively similar using word count in place of page count.

implementation of the reporting mandate. Whether this additional content helped to improve firms' information environment is an issue we examine empirically in section 6.

5. Construct validity tests

This section evaluates the validity of *StratScore* as a measure of strategy-related annual report commentary. Section 5.1 reports evidence based on a small sample manual scoring approach while section 5.2 presents results of large-sample correlation and causality tests.

5.1 Manual validation

We use a manual scoring exercise to provide evidence on the validity of our *StratScore* metric. Manual validation by domain experts is considered the most reliable means of assessing classification accuracy in the natural language processing literature. Our manual process involves the following steps. One co-author ($Author^{A}$) ranked the final sample of annual reports by *StratScore* and constructed three disclosure quality categories: High (top quartile); Low (bottom quartile); Medium (all remaining observations).[10] $Author^{A}$ then selected 20 reports from each category at random. The 60 reports were presented blind (i.e., no information on reports' classification and no details of the number of reports selected from each category) to two domain expert co-authors ($Author^{B}$ and $Author^{C}$). Each report was read and classified independently by $Author^{B}$ and $Author^{C}$ into one of the three quality categories based on an assessment of business model and strategy disclosures. $Author^{B}$ and $Author^{C}$ then returned their independent ratings to $Author^{A}$ for comparison against *StratScore* classifications.

[10] Consistent with the standard approach applied in natural language processing research, we use a three-level classification procedure to reflect the non-binary nature of disclosure quality and to reduce the unconditional probability of correctly classifying reports by chance.

Classifications by Author^{B} and Author^{C} agreed for 47 reports (78%). We use these 47 reports as the gold standard classification against which we assess *StratScore*'s ability to reproduce similar rankings.[11] Classification accuracy is assessed using measures of precision and recall (Manning and Schütze 1999). Precision measures the incidence of Type I errors and recall measures the incidence of Type II errors. Findings are presented in Table 3. The first row reports results for low quality classifications. Manual annotators independently classified 21 of 60 reports as containing low quality disclosures on business model and strategy. Sixteen of these 21 reports were drawn from the Low *StratScore* quartile, implying a recall rate of 76% (i.e., type II error rate equal to 5/21 or 24%). Two further reports originally drawn from the Low *Stratscore* quartile were not consistently rated as low quality by manual annotators, implying a precision rate of 89% (i.e., type I error rate equal to 2/18 or 11%). The same approach is used to compute precision and recall rates for Medium and High classifications in rows two and three, respectively. Precision and recall rates exceed 80% for High. Total precision and recall rates reported in the final row of Table 3 are well above 70% which is considered high for a three-way classification problem (Teufel et al. 2006). Findings indicate that *StratScore* correlates reliably with manual classifications of annual report commentary on business model and strategy.

5.2 Large-sample validity tests

Next we assess construct validity by correlating *StratScore* with factors predicted to explain variation in the extent of strategy-related annual report commentary, and by testing whether *StratScore* captures the exogenous increase in strategy reporting for LSE Main Market

[11] Disagreement between Author^{B} and Author^{C} on 13 reports highlights the inherent difficulty of scoring strategy-related disclosure. Nevertheless, inter-rater reliability scores for Cohen's Kappa and Krippendorff's Alpha are 0.67, which are respectable for a three-way classification problem (Teufel et al. 2006). Further analysis reveals that *StratScore* agrees with the classification assigned by one of the two domain experts in 11 of the 13 disagreement cases (85%), which provides further evidence that the metric outperforms a random classification benchmark.

firms as a consequence of the revision to the UK Corporate Governance Code 2010. Survey evidence confirms a significant increase in strategy-related reporting in response to the disclosure mandate, effective for fiscal years beginning on or after June 29, 2010 (Deloitte 2012). Critically, the rule did not apply to firms listed on the AIM section of the LSE, thereby affording a means of identification. If *Stratscore* provides a valid proxy for strategy-related annual report commentary then an incremental positive increase should be evident in Main Market firms' reports with fiscal year-ends from June 2011 onwards, relative to their AIM counterparts.

Further evidence regarding construct validity is provided by examining associations between *StratScore* and additional covariates that theory and practice predicts should correlate with the strategy-relevant annual report disclosure. However, the absence of a clear identification strategy for these variables means that standard concerns in the disclosure literature regarding endogeneity bias apply. Findings should therefore be interpreted with this caveat in mind.

Prior research highlights a series of firm-level characteristics associated with strategy-related disclosures. For example, organisationally complex firms are predicted to face more scope and greater demand to explain their strategic objectives and priorities (Gu and Li 2007). We therefore expect *StratScore* to be increasing in the degree of organisation complexity. The literature also finds a reliable positive association between disclosure levels and firm size. We expect this result to carry-over to strategy-related commentary for two reasons. First, large firms face greater incentives and pressure to explain and justify their activities to external stakeholders as a result of political cost considerations (Watts and Zimmerman 1986). Second, large firms tend to be characterized by greater operational complexity and as a consequence have more scope (and face a higher demand) to explain their objectives and approach to value creation. We therefore expect *StratScore* to correlate positively with firm size.

Financing constraints and the demand for external capital have also been linked with the provision of information on corporate objectives and business model. On the one hand, investors are expected to demand more strategy-related information to help them evaluate the investment proposition properly; on the other hand management are expected to face strong incentives to elucidate strategy in an effort to reduce information asymmetry and minimize incremental capital raising costs. Research confirms that management are more likely to discuss factors affecting their external environment when faced with financial constraints (D'Aveni and MacMillan 1990). We therefore expect *StratScore* to be increasing in the demand for external capital.

An externality of voluntary disclosure that is particularly pertinent to strategy-related commentary is proprietary costs arising from the information being exploited by competitors (Verrecchia 1983, Darrough 1993). The Department of Business, Skills and Innovation (DBIS) identified proprietary costs as the most important factor constraining U.K firms from disclosing more information on strategy, principal risks, and opportunities (DBIS 2010). We therefore expect *StratScore* to correlate negatively with competitive intensity.

Finally, the UK Corporate Governance Code (2010) was not the first attempt by UK regulators to encourage greater disclosure on strategy and approaches to value creation. As part of a revision to the 2006 Companies Act requiring *all* LSE-listed firms to include a business review in their annual report, the ASB issued a best practice statement on management commentary that encouraged directors to provide commentary on: corporate objectives and strategy; resources available to deliver those objectives; risks and uncertainties facing the entity; and factors likely to affect the firm's future development (ASB 2006). The best practice guidelines were effective for year-ends on or after March 31, 2006 and survey evidence suggests a significant fraction of UK-listed firms applied the guideline even though it was not mandatory

(PricewaterhouseCoopers 2007). Accordingly, we expect to observe a higher average value of *StratScore* for with year-ends on or after March 31, 2006.

We test our predictions using the following pooled OLS regression:

$$\begin{aligned} StratScore_{it} = {} & \alpha_0 + \alpha_1 CGC2010_{it} + \alpha_2 Main_{it} + \alpha_3 CGC2010 \times Main_{i1} \\ & + \alpha_4 ASB2006_{it} + \alpha_5 Segments_{it} + \alpha_6 Size_{it} + \alpha_7 Finance_{it} \\ & + \alpha_8 Competition_{it} + \sum_{k=1}^{K} \gamma_k Control_{kit} + \varepsilon_{it}, \end{aligned} \quad (3)$$

where variable definitions and associated predictions are as follows. *StratScore* is defined by equation (1). *CGC2010* is an indicator variable capturing changes in strategy-related reporting provisions in the UK Corporate Governance Code 2010 applicable to LSE Main Market firms' annual reports for fiscal years ending on or after June 30, 2011, and zero otherwise. *MAIN* is an indicator variable equal to one for the subset of firms listed on the LSE Main Market and zero for firms listed on AIM. Our primary construct validity test predicts an incremental increase in *StratScore* for LSE Main Market firms relative to their AIM counterparts in response to revisions to the U.K Corporate Governance Code 2010 ($\hat{\alpha}_3 > 0$). *ASB2006* is equal to one for fiscal year-ends on or after March 31, 2006 and zero otherwise, and captures any structural break in disclosure practices associated with the reporting guidelines issued by the ASB ($\hat{\alpha}_4 > 0$).[12] *Segments* is the number of operating business segments and is used to proxy for organisational complexity ($\hat{\alpha}_5 > 0$). *Size* is the natural logarithm of total assets and is used to proxy for firm size ($\hat{\alpha}_6 > 0$). *Finance* is an indicator variable equal to one for firms that either (a) display an ex ante demand for financing (the firm's free cash flow over the last three years divided by current assets is less than -0.5) or (b) raise finance during the fiscal year via either a seasoned equity issue or a

[12] We also present specifications including *ASB2006×Main* both for completeness and to ensure *CGC2010×Main* is not capturing effects associated with the 2006 regulatory change. We offer no formal prediction for *ASB2006×Main*, although survey evidence by PricewaterhouseCoopers (2007) suggests that Main Market firms were quicker to respond to the ABS (2006) guidelines than their AIM counterparts.

debt issue and it is used to proxy for financing constraints, and zero otherwise ($\hat{\alpha}_7 > 0$). *Competition* is the four-firm concentration ratio (multiplied by minus one) and is used proxy for competitive intensity within the sector ($\hat{\alpha}_8 < 0$).

Equation (3) also includes a vector of control variables. We include the total number of words in the financial statements section of the annual report as a control for overall corporate disclosure policy (*Wordcount_FS*). We also include a proxy for forward-looking commentary (*Forward*). While strategy-related commentary embeds forward-looking disclosures about corporate priorities and trends expected to affect future cash flows (Gu and Li 2007, FRC 2014), reporting on strategy involves a broader focus. We therefore include the number of forward-looking words from El-Haj et al. (2018) to ensure *StratScore* is capturing more than just the presence of future-orientated statements. Extant research also highlights poor performance and growth opportunities as potential drivers of strategy-related disclosures, although theory and empirical evidence yields ambiguous predictions regarding the sign of the associations (Diffenbach and Higgins 1987, Gu and Li 2007, Lu and Tucker 2012). We therefore include return on assets (*ROA*) and the book-to-market ratio (*BM*) as controls for performance and growth opportunities, respectively, but offer no directional predictions. We also include the lagged value of the dependent variable because strategy and disclosure thereon are likely to be sticky over time. Finally, we estimate equation (3) inclusive of industry and time fixed effects. Summary statistics for test and control variables are reported in Panel B of Table 2.

Figure 1 plots sample mean values for *StratScore* by year and provides qualitative insights concerning our main prediction. The *StratScore* series for both Main Market and AIM firms are upward sloping, indicating that the provision of strategy-related commentary increased for both groups across the sample period. However, the slope for Main Market firms is

noticeably steeper. ASB (2006) guidelines implemented for fiscal years ending on or after March 31, 2006 appear to have triggered divergence in strategy reporting practices even though the guidelines applied to all listed firms. A more significant divergence in *StratScore* levels is evident following the 2010 revision to the UK Corporate Governance Code. The incremental increasing trend in *StratScore* post-30 June 2011 for Main Market firms is consistent with our score capturing regulation-driven changes in strategy-focused commentary.[13] Figure 1 also supports the DID parallel paths assumption, with less dramatic divergence in strategy reporting practices between the two groups evident during the four years prior to June 30, 2011.

Summary statistics and coefficient estimates for equation (3) are presented in Table 4. Adjusted-R square values range from a high of 74% in Model 4 where we control for the lagged value of the dependent variable to a low of 28% in Model 6 where we replace *StratScore* with a scaled version of the measure that controls for an entity's general disclosure propensity. These findings suggest that the predicted drivers of strategy-related disclosure in equation (3) are collectively able to explain a substantial amount of the variation in *StratScore*.

Model 1 confirms a significant increase in the average level of strategy-focused disclosure following implementation of the UK Corporate Governance Code 2010. Model 2 extends the regression to include the *CGC2010×Main* interaction and as predicted the estimated coefficient loads positive and significant, confirming that *StratScore* detects the incremental increase in strategy-related annual report commentary mandated for Main Market firms by the 2010 revision to the UK Corporate Governance Code. The result is incremental to controls for

[13] Figure 1 suggests the presence of learning effects for strategy-related reporting. Although we offer no formal prediction regarding such effects, evidence of incremental improvement in the quality of strategy-related commentary has been heighted in the professional literature (PwC 2016). An alternative explanation for the deferred spike in *StratScore* is an increase real economy effects involving strategic implications such as growth in merger and acquisitions. Data from the IMAA Institute do not support this explanation (https://imaa-institute.org/m-and-a-uk-united-kingdom/).

overall annual report disclosure policy (*Wordcount_FS*). The positive coefficient on *CGC2010×Main* is also robust to controlling for reporting changes associated with the ASB guidelines introduced in 2006 (Model 3), and inclusion of the lagged value of *StratScore* (Model 4). *CGC2010×Main* continues to load in Model 5 after including a proxy for forward-looking disclosure, confirming that *StratScore* captures aspects of strategy-related disclosure incremental to and distinct from forward-looking commentary per se. Model 6 includes the text-based proxy for competition proposed by Li et al. (2013). The coefficient on *CGC2010×Main* remains positive, consistent with descriptions of business model and strategy extending beyond a discussion of competition and *StratScore* reflecting this broader focus. Finally, Model 7 presents results using scaled versions of *StratScore* and forward-looking disclosure.[14] The coefficient estimate on *CGC2010×Main* remains positive and significant at the 0.01 level.

Results for additional test variables in equation (3) are also consistent with predictions. *Segments*, *Size*, *Finance* and *Competition* load with their expected coefficient signs and only *Competition* in Models 1-3 and 6 is not significant at conventional levels. *ASB2006* loads positively in Models 1 and 2 reflecting introduction of the business review guidelines in 2006 and the specific recommendation for more strategy-related commentary (ASB 2006).[15] Collectively, findings presented in Table 4 suggest that *StratScore* varies in manner consistent with theory and regulatory guidelines concerning strategy-related disclosure.

Although findings in Table 4 are robust to controlling for *Wordcount_FS* and using scaled *StratScore*, it is still possible that the results reflect overall disclosure policy rather than strategy-

[14] We use the number of words in the financial statements section of the report rather than the number of words in the narratives section because the latter includes aspects of strategy-related commentary present in *StratScore*. In contrast, commentary in the financial statements component of the report provides a firm-level measure of disclosure propensity that is largely independent of management's specific reporting policy on strategy.

[15] Results for Models 3 and 4 also support survey evidence from PricewaterhouseCoopers (2007) suggesting that traction on the ASB's guidelines was limited primarily to Main Market firms: the *ASB2006* main effect term capturing average reporting behaviour by AIM firms is negative while (*ASB2006* + *ASB2006×Main*) is positive.

specific reporting. To further assess whether *StratScore* captures a dimension of disclosure policy distinct from the general trend towards longer annual report narratives (EY 2015), we conduct a placebo test where *StratScore* in equation (3) is replaced with measures of aggregate annual report narrative disclosure. If results in Table 4 reflect the general increase in disclosures unrelated to strategy then *CGC2010×Main* should continue to load using broader measures of annual report narrative commentary. We use two proxies for general report-level disclosure policy: the number of words in the financial statements component of the annual report (see footnote 16) and the residual word count for the narrative component of the annual report after excluding the following three sections where manual inspection suggests strategy-related commentary is most often located: summary highlights, the letter from the board chair, and management commentary. (This second disclosure proxy provides a particularly tough placebo test for *StratScore* because the residual component of the narratives almost certainly contains some strategy-related commentary for some firms.) Findings using both measures are reported in Table 5. The coefficient estimate on *CGC2010×Main* is indistinguishable from zero in both models. Findings suggest that results presented in Table 4 using *StratScore* are indeed capturing strategy-specific annual report disclosures rather than broader disclosure policy. Meanwhile, evidence that *Segments*, *Size* and *Financing* continue to load positively in one or both models supports the view that these characteristics are associated with overall corporate disclosure policy as well as strategy-specific reporting and as such cannot be relied on to provide a definitive construct validity test for *StratScore*.

Results reported in Tables 4 and 5 confirm *StratScore* as a valid measure of strategy-related annual report commentary. Critically, findings suggest that *StratScore* captures disclosures that are distinct from forward-looking commentary, competition, and general trends in the volume of annual report content. Evidence also indicates that the structural shift in

narrative reporting policy following revision to the UK Corporate Governance Code 2010 centred on strategy-related content rather than annual report commentary more generally.

6. Capital market consequences of enhanced strategy-related disclosure

6.1 Investor uncertainty

Evidence presented in the previous section reveals a structural increase in strategy-related annual report commentary among LSE Main Market firms for fiscal year-ends on or after June 2011 in response to the 2010 revision to UK Corporate Governance Code. In this section we exploit the exogenous change in strategy-related reporting for a subset of LSE firms to construct a DID test of the capital market consequences of enhanced annual report discussion of the value creation process.[16]

Consistent with the view that annual report strategy-related commentary establishes a broader information framework that helps users resolve uncertainty about firms' value creation process, we test whether increase in strategy-related disclosure reduce information asymmetry in the form of dispersion in analysts' earnings forecasts, the magnitude of forecast errors on earnings and target price, and the bid-ask spread (Lehavy et al. 2011, Hail and Leuz 2006). We include analyst-based measures because although analysts likely had private information on industry competition, business models and company strategy, survey evidence reported by PwC (2006a: 4-5, 2006b: 18) indicates their knowledge was far from complete.

We use the following OLS regression to evaluate the capital market consequences of enhanced reporting on strategy and business model:

[16] Figure 1 confirms the validity of the equality of pre-treatment trends assumption. While the operating and financial reporting guidelines issued by the ASB (2006) also marked an increase in strategic commentary, the policy innovation offers a less promising identification strategy because application was voluntary and the guidelines applied to all LSE firms.

$$CMC_{it}^{m}=\gamma_0+\gamma_1 CGC2010_{it}+\gamma_2 Main_{it}+\gamma_3 CGC2010\times Main_{it}+\sum_{p=1}^{P}\lambda_p Controls_{pit}+v_{it} \quad (4)$$

The dependent variable is a proxy for capital market consequences as measured by information asymmetry for firm *i* in fiscal year *t*+1(*m* equals earnings forecast dispersion, absolute earnings forecast error, absolute price forecast error, or bid-ask spread). Analyst forecast dispersion is the standard deviation of individual analyst earnings forecasts from I/B/E/S for year *t+1* issued during the period between the earnings announcements for year's *t* and *t+1*. Absolute forecast errors on earnings are equal to the absolute value of the difference between I/B/E/S actual earnings for year *t*+1 and the initial median I/B/E/S consensus earnings forecast outstanding for year *t+1* following the announcement of earnings for year *t*. Absolute forecast errors on target prices for year *t* are equal to the absolute value of the difference between the target price following the earning announcement for year *t* and the 12-month-ahead actual stock price (Bilinski et al., 2013). Bid-ask spread is the rolling average of the monthly spread (ask minus bid price divided by the average of the bid and ask price) from Datastream computed over the fiscal year *t+1*. All four information asymmetry proxies are scaled by lagged stock price. *CGC2010* and *Main* are defined as in equation (3) and our DID estimator is $\hat{\gamma}_3$, which captures the incremental change in information imperfections for LSE Main Market firms relative to their AIM counterparts in response to the UK Corporate Governance Code 2010 strategy reporting mandate. We interpret $\hat{\gamma}_3 < 0$ as evidence that enhanced strategy-related reporting leads to a reduction in information asymmetry. Since compliance with the strategy mandate gives firms the option of explaining non-disclosure, the perfect compliance assumption underpinning the DID approach (Bundell and Dias, 2009) may be violated in our setting, in which case $\hat{\gamma}_3$ estimates treatment effects for the marginal discloser only (i.e., the marginal treatment effect). In reality,

however, few (if any) Main Market firms exercise the non-disclosure option. Finally, *Controls* is a vector of *P* covariates comprising the *CGC2006* indicator variable and its associated two-way interaction with *Main*, factors identified by prior research as effecting information asymmetry in the form of firm size (*Size*), growth options (*BM*) and operational complexity (*Segments*), and properties of annual report narratives in the form of disclosure propensity (*WordCount_FS*), forward-looking commentary (*Forward*), and annual report readability proxied by the Fog index (*Fog*) (Li, 2008).

The initial sample of 9,127 observations is restricted to firms with at least two observations in both *CGC2010* subperiods. Final sample sizes for each version of equation (4) are determined by available observations for the m^{th} *CMC* proxy. Sample sizes range from 3,399 where *m* equals absolute forecast error on price, to 6,027 observations where *m* equals the bid ask spread. All *CMC* proxies are winsorized at their extreme 0.5 percentiles and log transformed.

Table 6 reports results for equation (4) estimated using the four proxies for information asymmetry. The DID estimator on *CGC2010*×*Main* loads negative for all proxies in Table 6, although results for absolute forecast errors on price are not statistically significant at conventional levels. Estimated effects on most CMC proxies are also substantively significant. For example, the post-2010 decline in price-scaled forecast dispersion is 51 percent larger for Main Market firms relative to AIM firms after controlling for all other factors in the model.[17] Comparable effects for absolute price-scaled forecast error (earnings) and price-scaled bid-ask spread. Collectively, these findings provide initial evidence that the strategy-related disclosure mandate introduced in the 2010 revision of the UK Corporate Governance Code improved the disclosure environment for LSE Main Market firms relative to their AIM counterparts.

[17] Dependent variables in equation (4) are log transformed and therefore coefficient estimates indicate the percentage change in the dependent variable for a one-unit (category) change in a given explanatory variable.

Our DID design relies on the untreated sample of AIM firms being otherwise similar to our treated sample of Main Market firms. AIM constituents face the same economic conditions and financial reporting rules to their Main Market counterparts. Nevertheless, the average AIM firm is smaller, more risky, followed by fewer analysts, and faces lower stock market liquidity relative to the typical Main Market constituent. We use the synthetic control sample method to assess the robustness of our findings to control firm selection (Abadie and Gardeazabal, 2003). As this method applies to a strictly balanced sample, we restrict our information asymmetry proxy to bid-ask to maximize sample size. We begin by constructing a balanced sample with available data over six years surrounding the revision of the UK Corporate Governance Code in 2010. Following Acemoglu et al. (2016), we match each LSE Main Market (i.e., treated) firm in the first year under treatment (i.e. first year where the revision of the code is effective, fiscal year 2011) to a combination of AIM (i.e., untreated) firms that closely match the treated unit over the pre-treatment period fiscal years 2008-2010. Matching is performed using a convex weighting matrix that minimizes the Euclidean differences in bid-ask spread on each of the pre-treatment periods. Outcomes for the resulting synthetic control sample of AIM firms are then projected into fiscal years 2011-2013 using the weights identified by the pre-treatment comparison. Projected outcomes for the synthetic control sample represent the counterfactual for the treated unit.

Findings are presented graphically in Figure 2. Differences in average bid-ask spread between Main Market and synthetic control groups are not significant by construction before fiscal year 2011. A sharp divergence in information asymmetry between Main Market and AIM firms is evident following implementation of the strategy reporting mandate: log spreads drop sharply in the first and second treatment years for Main Market firms and remain low thereafter, whereas log spreads exhibit a minor drop in the first treatment period followed by a steady rise in

the following two years for the synthetic control sample. Conclusions are therefore consistent with the DID results reported in Table 6.

Although results in Table 6 are consistent with enhanced reporting requirements on strategy reducing information asymmetry, it is possible that other unmodeled changes in governance and reporting practices occurring after 2010 for Main Market firms may be driving the results. We address this concern using two approaches. First we re-estimate equation (4) after replacing *CGC2010×Main* with separate interactions on *Main* for each sample year. Untabulated results provide no evidence of an alternative structural break between July 2011 (when the first annual reports implementing the new reporting rules were published) and the end of the sample period. Further, none of the year interactions load negatively before July 2011. The evidence does not support the presence of an alternative cause for the results reported in Table 6 that we attribute to the change in strategy-focused reporting.

Our second approach to assessing the validity of inferences drawn from Table 6 involves estimating an alternative specification of equation (4) where the indicator variable *Main* and its two-way interaction with *CGC2010* are replaced with a proxy capturing the subset of firms with the largest improvement in *StratScore* from the pre- to the post-mandate period. Using the restricted sample of firms with at least two observations in both periods, we construct an indicator variable (*StratScore_increase*) equal to one for firms with a top quartile difference between average *StratScore* computed using pre-mandate observations and average *StratScore* computed using post-mandate observations, and zero otherwise. We also include firm fixed effects to account for the endogeneity of the rise in *StratScore*. Results are reported in Table 7. Consistent with the findings presented in Table 6, *StratScore_increase* loads negatively ($p < 0.05$) in all models. Further, coefficient estimates imply economically significant reductions in information asymmetry ranging from 30 percent for forecast errors on price to 73 percent for bid-

ask spread. Evidence of a direct link between improvements in firms' information environment and increases in their *StratScore* value surrounding the 2010 revision to the UK Corporate Governance Code helps alleviate concern that our DID tests reported in Table 6 are capturing the effect of other regulatory or economic shocks unrelated to changes in strategy-related commentary. We nevertheless acknowledge that these tests cannot rule out completely the possibility that the observed reduction in investor uncertainty reflects unspecified reporting and governance innovations co-occurring with the 2010 strategy disclosure mandate.

6.2 Price discovery rate

If strategy-focused management commentary provides a contextualizing framework that helps investors better understand the consequences of realized financial and non-financial performance for future cash flows then such disclosure may improve the speed with which information is incorporated into stock prices following publication of the annual report. Callen et al. (2013) discuss how investors' ability to interpret new information as it arrives is conditioned by their pre-existing information set. A clear grasp of an entity's strategic objectives and business model for delivering those objectives arguably provides the core conditioning framework necessary to understand the mapping from short-term activities and performance into long-term value creation (FRC 2014). Building on Callen et al. (2013), greater transparency about corporate strategy should facilitate an improved appreciation of the value creation process and lower uncertainty with respect to valuation parameters, as reflected in the speed at which price converges to fundamentals. We therefore test for a reduction in the stock price delay in response to the structural increase in strategy-related annual report commentary resulting from the 2010 revision to the UK Corporate Governance Code. Following Hou and Moskowitz (2005) and Callen et al. (2013), we approximate the average delay with which information is impounded into

stock prices following publication of the annual report using the difference between the explanatory power of firm-specific regressions of weekly stock returns on (a) contemporaneous and four-lagged weekly market returns and (b) contemporaneous market returns only, estimated over the 12-month period following the annual report release. We also construct a firm-specific delay measure computed as the difference between the explanatory power of firm-specific regressions of weekly stock returns on (a) contemporaneous market and four-lagged firm-specific weekly returns and (b) contemporaneous market returns only.

Table 8 reports results for equation (4) with the CRM proxy equal to stock price delay over the 12-month window following the annual report release. For consistency with other CRM tests we use log transformed stock price delay measures. Results for the DID model are presented in columns 2 and 3. The DID estimator loads negatively in both models as expected, although statistical significance is marginal ($p < 0.1$). A similar pattern is evident in the final two columns in Table 8 for models estimated with the *StratScore_increase* indicator from Table 7 and firm fixed effects. The estimated coefficient on *StratScore_increase* is negative in both models. However, statistical significance is marginal and sensitive to the choice of delay proxy. Note also that the coefficient estimate on *CGC2010* capturing the structural shift for AIM firms is positive in all four models, suggesting that stock price delay increased for AIM firms while remaining broadly constant for Main Market firms. Results therefore provide weak statistical support consistent with enhanced strategy-focused commentary lowering investor uncertainty through more rapid assimilation of annual report contents into stock price.

6.3 Cross-sectional analysis

Our final set of analyses tests whether the information benefits documented in Tables 7 and 8 vary cross-sectionally with the quality of firms' information environment prior to the

introduction of the strategy reporting mandate. Specifically, we test if the reduction in information asymmetry in response to the increase in strategy-focused annual report disclosure is more pronounced for firms with poor information environments and low strategic transparency in the pre-2010 period, as proxied by low analyst coverage. Analysts generate information about strategy and business model through their forecasting and valuation activities. For firms with low analyst coverage, externally-generated information on strategic goals and models for long-term value creation is likely to be more limited relative to their more intensely followed counterparts. In such cases, management commentary on strategy and business model is expected to provide a particularly valuable source of information for investors. We therefore test whether reductions in information asymmetry in response to increases in strategy-focused annual report commentary are more marked for low coverage firms prior to the regulatory change.

We extend findings reported in Tables 7 and 8 by allowing the coefficient estimate on *StratScore_increase* to vary with the level of analyst coverage. We create an indicator variable (*LowCover*) equal to one for firms in the lowest quartile of the pre-2010 distribution of analyst following and then test for a negative coefficient estimate on *StratScore_increase*×*LowCover*. Results reported in Table 9 confirm that the information benefits to enhanced strategy-focused commentary are indeed more pronounced for the low analyst coverage group. As expected, *StratScore_increase*×*LowCover* loads negatively at the one percent significance level in the majority of cases.[18] Bid-ask spread is the only information proxy in Table 9 where the information gains to enhanced strategy reporting are independent of analyst following. In contrast, where the dependent variable is absolute forecast error on earnings and price, the incremental coefficient on *StratScore_increase*×*LowCover* is more than three times larger in

[18] We are unable to present results for forecast dispersion in Table 9 because the variable is indeterminate for observations where analyst coverage is very low.

magnitude than the corresponding main effect estimate for moderate and high coverage firms; and for price delay models reported in the final two columns of Table 9, significant information gains are exclusive to low coverage firms.

7. Summary and conclusions

We develop an objective, replicable measure of strategy-related commentary that we use to explore the benefits of enhanced annual report disclosure on strategy and business model. Our measure is based on a customized list of strategy-related keywords and phrases drawn from the extant strategy literature. To reduce risk of Type I errors we also use an objectively identified corpus of strategy commentary to construct a weight for each element in our list that reflects the conditional probability that it is used in a strategy context. Construct validity tests confirm our measure is a legitimate proxy with which to evaluate the capital market impact of an increase in strategy-related annual report disclosure. We therefore proceed to test whether enhanced strategy-focused disclosure reduces investor uncertainty proxied by analyst forecast dispersion, analyst absolute forecast errors on earnings and price, and bid-ask spread. We also test whether the contextualizing role of strategy-related commentary increases the speed of price discovery over the 12-month period following publication of the annual report. Our identification strategy exploits the 2010 revision to the UK Corporate Governance Code requiring firms with a primary listing on the LSE (but not their AIM counterparts) to explain their strategy and business model in their annual report. DID results reveal a statistically and economically significant incremental decrease in investor uncertainty for LSE primary listing firms in response to the disclosure mandate. We also find evidence of a decrease in stock price delay that is particularly pronounced for listed firms with low analyst coverage.

Two caveats are appropriate when interpreting these results. First, the perfect compliance assumption underpinning our DID analysis is almost certainly violated and as such our approach estimates treatment effects for the marginal discloser only (i.e., the marginal treatment effect). Caution is therefore advised when attempting to generalize treatment effects beyond this group. Second, it is possible that other unobserved regulatory changes to governance or reporting occurring during the sample window may be responsible for reported improvements in firms' information environment. We address this issue by estimating models using the change in strategy-related commentary. Results are broadly consistent with our DID analysis and as such provide comfort that our findings reflect changes in the provision of strategy-related annual report commentary, although we acknowledge that we cannot rule out other effects entirely.

Theory and practice stress the potential benefits to capital market participants associated with a clear understanding of firms' objectives and their approach to creating and preserving value. In response, best practice reporting guidelines increasingly promote disclosures on strategy and business model. A small body of work examines the causes and consequences of voluntary strategic plan disclosures outside the annual report. However, despite increasingly loud calls from regulators and investors for management to provide more strategy-focused commentary in their annual reports, the capital market impact of such disclosures has not been examined. Our analysis speaks to this question and provides the first systematic evidence that providing strategy-focused annual report commentary can help improve the quality of firms' information environment. The findings are timely given the SEC's request for views on mandating disclosure of registrants' business model in the MD&A (SEC 2016: 60).

Appendix A: Procedure to construct business model and strategy wordlist

This appendix describes the procedure used to identify the 709 words and phrases (hereinafter referred to as n-grams, where n is ≥ 1) used as the basis for our *StratScore* measure of management commentary on strategy and business model. The basis for our strategy-focused list of n-grams is all elements included in the indexes of the following five leading textbooks on business strategy: Porter (1985), Barney and Clark (2007), Rumelt (2011), Magretta (2012), and Grant (2013). For completeness, we supplement this list with the 517 strategy-related key terms reported by Ronda-Pupol and Guerras-Martin (2012, Appendices 1a through 1c) in their corpus analysis of the evolution of the concept of strategy in the academic literature.

Pooling the contents from these six strategy sources yields an initial unedited list of 4,384 n-grams. The resulting list was then curated manually by members of the research via three rounds of edits aimed at minimizing the Type I error rate and maximizing parsimony. The first editing round applied the following two exclusion criteria:

- Proper nouns (e.g., names of firms, authors, managers, countries, sectors, products, etc.) unrelated to strategy, which if retained will introduce Type I errors in our frequency counts;
- Duplicate use of the same n-gram by one or more authors.

Applying these criteria resulted in the exclusion of 801 n-grams. The refined list of 3,584 distinct n-grams is available in an Excel worksheet titled "master list of 3,584 original unique phrases" (hereinafter "master list") via an online appendix (DOI: 10.17635/lancaster/researchdata/232). The worksheet also includes the original source(s) of each n-gram.

The second round of edits involved three members of the research team working collectively to identify further candidates for exclusion based on the following two criteria:

- Terms that are either generic and unlikely to be linked directly to strategy (e.g., "economics", "governance", "investments", etc.) or sufficiently general such that while they may occur in strategy-related commentary they are also likely to appear frequently in non-strategy-focused disclosures (e.g., "profit", "asset", "cost", etc.). Retaining such terms will result in a high Type I error rate. These exclusions are marked in yellow in the master list Excel worksheet;
- Complex or idiosyncratic terms that are so unique that the processing cost of retained the term for regular expression comparisons outweighs the probability of the term appearing in normal written English. Examples include "offensive strategy competitor interrelationship" (Porter 1985), "joint venture agreements and strong form trustworthiness" (Barney and Clark 2007), and "advantage and silver machine problem" (Rumelt 2011). These exclusions are marked in green in the master list Excel worksheet.

N-grams that survived this second editing round are marked in red in the master list Excel worksheet. Finally, the third editing sought to simplify the interim wordlist. Specifically, we applied the following *exclusion* criteria aimed at identifying shorter n-grams that captured an aspect of strategy and which formed part of at least one of the phrases in the interim list:

- Redundancy due to inclusion of multiple versions of the same underlying construct. For example, the index for Grant (2013) includes the following terms all relating to core strategic concept of bargaining power: "bargaining power bilateral monopolies", "bargaining power cost advantage", and "bargaining power internationalization". In this case we replaced multiple redundant phrases with the single n-gram "bargaining power";
- Ambiguities associated with acronyms and case-sensitive representations that require the addition of disambiguation conditions. For example, "aim" may be related to strategy commentary whereas the acronym for Alternative Investment Market is not strategy-related.

Similarly, “international” may feature in strategy commentary whereas “International” is more likely to be part of a firm name where it does not occur at the start of a sentence;

In addition, we also applied *inclusion* criterion designed to incorporate both English and American English spelling (e.g., “centralise” and “centralize”) and to ensure complete coverage of alternative versions of multiword expressions. For example, while the initial master list includes several representations of research and development (e.g., “R and D”, “Research and Development” and “research and development”), other feasible presentations such as “R&D” and “Research & Development” would be excluded unless the expressions are added manually. Similarly, while the initial mater list includes “switching costs” (Grant, 2013, Margretta 2012, Porter 1985), the natural alternative representation “cost of switching” would be excluded without this manual intervention step. Finally and where appropriate, we applied lemmatization to selected elements to ensure various inflections (including plurals) are captured.

This third round of edits involved two co-authors applying inclusion criteria to each n-gram independently, with disagreements resolved by a third co-author. Retained n-grams were then considered independently by two co-authors for inflecting or lemmatizing, with a third co-author again serving as judge where disagreements occurred. This refining and cleaning process occurred over multiple iterations resulting in a final list of 709 n-grams (including inflections, alternative spellings, and alternative representations of the same term or phrase).

The final list of 709 n-grams is presented in Table A1, partitioned between the 231 n-grams with a positive weight in the calculation of *StratScore* (Panel A) and the remaining 478 n-grams that receive a zero weight in the *StratScore* calculation (Panel B). The list is also available via an online appendix that provides a full reconciliation to the 3,584 n-grams in the original master list (doi 10.17635/lancaster/researchdata/232).

Table A1 List of 709 strategy-focused n-grams used to compute *StratScore*. N-grams preceded by (L), (U) and (LU) are lower case only, upper case only, and lower or upper case, respectively. # indicates lemmatizing.

Panel A: 231 n-grams with a non-zero weighting in final *StratScore* measure

Abilities
Ability
adapt#
advantage#
Aim
Aiming
Aims
ambition#
analytic# tool#
Attack
attack#
Attractive
attractiveness
balanced scorecard
barrier# to entry
behavior#
behaviour#
benchmark#
best practice#
bottom up
bottom-up
brand#
Bundling
business model#
business process#
business scope
business system#
capabilit#
centraliz#
centre# of excellence
cluster#
co operation
co ordinat#
cohesiveness
collaborat#
commercialis#
commodity product#
compete
competenc#
competing
competitive
competitive advantage#
competitive force#
competitor#
complement#
complexity
containment
coordinat#
core competenc#
corporate portfolio
cost driver#
cost improvement#
cost# of switching
create#
creati#
cultural
culture
customer#
customiz#
decentralis#
decentraliz#
defensive
demand-side
demographic#
deploy
differentiat#
distribution channel#
divers#
driver
drivers
economies of scale
efficien#
elastic
emerging
employee#
employment cost#
enhanc#
entrepreneur#
entry
environmental factor#
equilibrium
evolv#
exclusive
experience curve
fail#
failure
firm level
first mover
first-mover
flexibility
focus#
follower#
fragment#
game chang#
game-chang#
geographic
global#
goal#
growth
hands-on management
heterogeneity
holistic
human capital
incubate
incubating
inflection point#
innovate#
innovation#
innovative
innovator#
intellectual property
internal environment#
internal organisation
internalis#
internationalisation
interrelate#
invent
invest
investing
job satisfaction
know-how
knowledge
KPI
lead time#
leader#
lead-time#
legal constraints
leveraged
leveraging
life cycle
life-cycle
local#
long term
long-term
low cost
low-cost
M & A
M&A
managers
markets
maximise profit#
measure# of profit
merger# & acquisition#
merger# and acquisition#
mission
multi-business#
multiple
network effect#
network of
next big thing
norms
objectiv#
operat# effectiveness
operations management
opportunism
organisational structure
our network#
out sourc#
out-sourc#
performance criteria
performance indicator#
performance measure#
performance monitoring
performance target#
pioneering
plant size
political factor#
portfolio management
predatory
principles
procurement
profit measure#
profit pool#
property rights
proprietary
purchasing asset#
purchasing power
qualification cost#
quality
range of business#
reliance on
rely on
reputation#
resource-
respond
responding
responsiveness
risk#
roadmap
scalable
scale
scope
scorecard
seller#
service
shared
shareholder
sharing
shift#
shifting
skill#
social factor#
social policy
social responsibility
socially responsible
specializ#
stakeholder#
standardis#
strateg#
stream# of profit#
superior
supply-side
sustain#
switch back
switching back
synerg#

target#
team-based
technical#
technolog#
trade off
uncertain#
unique
uniqueness
utilise
utilization
utilize
value chain
value#
vision

Panel B: 478 n-grams with zero weighting in final *StratScore* measure

acquire
acquisition#
activit#
administrat# costs
administrat# distance
alliance#
ambitious
analysis
analytic# framework
anti business
anti compet#
anti trust
antibusiness
anti-business
anticompet#
anti-compet#
antitrust
anti-trust
armï¿½s length contract#
armï¿½s-length contract#
arms length contract#
arms-length contract#
asset assignment
asset# purchase#
assign# asset#
assignment of asset#
attain#
attitude survey#
bargaining power
barrier# to exit
barrier# to imitation
barrier# to mobility
beachhead#
blocking position#
blueprint
boundar#
bundle
bundled
business identity#
business linkage#
business unit#
business-process#
buyer cost
buyer performance
buyer power
buyer purchase criteria
buyer segment#
headwind#
heterogeneous
hierarch#
horizontal organisation
buyer type#
buyer value
buyer# need#
buyer# perception#
buyer/supplier
buyer-supplier
capacit#
capital cost#
carried out
carry out
carrying out
causal ambiguity
center# of excellence
centralis#
centralised direction
centralised redirection
centralised-direction
centralised-redirection
centralized direction
centralized redirection
centralized-direction
centralized-redirection
chain link
chain of value
chain-link
challeng#
chang#
cherry pick#
cherry-pick#
co operate
co operating
coalign#
co-align#
coalition#
cognitive factor#
cognitive map#
coherence
coherent
cohesion
cohesive
commercializ#
comparative advantage#
comparative disadvantag#
competition
competitive disadvantage#
complex
concentration
licensing
location
locational difference#
longer term
condition of duplication
configur#
conflict#
conglomerate
consolidate
consolidating
consolidation
continuity
control of timing
conventional wisdom
conventional-wisdom
cooperate
co-operate
cooperating
co-operating
cooperation
co-operation
co-ordinat#
corner solution
cost advantage#
cost disadvantage#
cost leader#
cost leading
cost# of sharing
crisis
cross licens#
cross subsidi#
cross-licens#
cross-subsidi#
customis#
decision#
demand
deployment#
deregulation
design
deterren#
develop#
diffusion
disadvantage#
disinvest#
dispos#
disrupt#
dissonance
distribution of assets
divest#
division of labor
division of labour
organizational identity
organizational intelligence
organizational rigidity
organizational routinization
dominant
dominate
downstream channel#
downturn
dynamic
economic#
economies of scope
economy of scale
economy of scope
emergent
emotional climate
emotional intelligence
enabl#
engineering overreach
enter
entered
entering
entrant
entry barrier#
envelopment manoeuvre
evolution#
exclusiveness
exclusivity
execut#
exit barrier#
exit#
exogenous
experience goods
exploit#
external
external environment#
financial interrelat#
financial management system#
firm specific
firm-level
firm-specific
fit
force#
franchisee
franchiser
franchising
franchisor
functional
geographical
hands off management
hands on management
hands-off management
quantity
R & D
R and D
R&D

horizontal organization
horizontal structure
horizontal system#
hostile takeover#
HR
imitability
imitat#
imitation barrier#
immobility barrier#
impact on profit#
implement#
incubator
incumbent
indivisibilit#
industrial buyer#
industries
industry
inelastic
inertia
information system#
information technolog#
infrastructure
input cost#
institutional factor#
intangibl#
integrat#
intend
intent
interfirm network#
inter-firm network#
internal force#
internal organization
internaliz#
international
internationalization
interrelationship#
invested
investment in
investments in
invisible asset#
IP
IPR
isolating mechanism
IT
joint ventur#
judgement
leading
leadtime#
lean production
leapfrog#
leap-frog#
learning
learning-by-doing
separation
series of actions
shape
shaped
Shaping
longer-term
loss of coherence
lower cost
lower-cost
macroscenario#
macro-scenario#
make or buy
make-or-buy
management control
management system#
management team#
managerial talent
mapping
market
marketing costs
marketing spend
matrix organisatio#
matrix organizatio#
maximize profit#
milestone#
mobility barrier#
monopolies
monopoly
mov# to action
multibusiness#
multifunctional units
multi-functional units
national
need# of buyer#
network externalit#
new entrant#
next big-thing
obscuring profit
obstacl#
operat# relatedness
operating organisation#
operating organization#
opportunit#
organisation# redesign
organisational analysis
organisational decisions
organisational design
organisational ecology
organisational identity
organisational intelligence
organisational rigidity
organisational routinisation
organisational slack
organization# redesign
organizational analysis
organizational decisions
organizational design
organizational ecology
standards
Static
stealth marketing
strength#
structural analysis
organizational slack
organizational structure
outsourc#
patent#
path dependen#
penetrate
penetration
perception# of buyer#
performance analysis
performance control
performance diagnosis
performance incentive#
personnel rotation
pioneer
pioneered
plan
plann#
plans
political distance
position
positive feedback
power of buyers
preempt#
pre-empt#
price cut#
price discrimination
price movement#
price premium
price sensitive#
pricing
privileged information
process#
product#
product-line
profit maximis#
profit maximiz#
profit source#
profit stream#
profitab#
program# of action
promotion from within
proposal#
protection of turf
proximate objective#
psychographic#
psychological contract
psychological factor#
purchase criteri#
purchase occasion
purchase power
purchase# of asset#
purchased input cost#
pursue
supplier
supply
survey of attitude#
surviv#
switch cost#
real option#
recession#
reciprocal dealing
reconfigur#
reengineer#
re-engineer#
refutation
regime of appropriability
region#
relational contract#
relationship
relative cost
repetitive
reposition#
research
research & development
research and development
resource#
response
restructur#
retaliate
retaliation
revolutionary
revolutionise
revolutionize
rival#
road map
road-map
routinis#
routiniz#
rules for riches
salience effect#
satisfactory
satisfic#
scale economies
scale economy
scenario analysis
scenarios
schumpeterian shock#
scope economies
scope economy
secto#
segment#
self organisation
self organise
self organization
self organize
self-organisation
self-organise
self-organization
self-organize
sense of purpose
sentiment
transfer#
trust#
type of buyer
utilisation
utilised

signal#	structural break#	switch-back	utilising
social equity	stuck in the middle	switched back	utilized
social herding	stuck-in-the-middle	switching cost#	utilizing
social legitimacy	sub technolog	swot	weaken#
social system#	substitut#	tail-risk#	weakness#
source# of profit	sub-technolog	takeover#	win
sourcing choice#	subtechnolog#	task#	winner
specialis#	succeed#	team based	winning
Spillover	success factor#	team production	wins
standalone influence	success#	threat#	
standardiz#	superiority	trade-off	

Our starting point for the list is the indexes of following five leading business strategy textbooks: Porter (1985), Barney and Clark (2007), Rumelt (2011), Magretta (2012) and Grant (2013). For completeness, we also include n-grams reported by Ronda-Pupol and Guerras-Martin (2012, Appendices 1a through 1c). The initial pooled list comprised 4,384 words and phrases, which reduced to 3,584 following exclusion of proper nouns (firms, products and authors) and other irrelevant content. The resulting list was then curated manually and independently by two members of the research team to: (a) remove generic words such as "profit", "industry", and "asset" that are unlikely to discriminate between strategy-specific content and other aspects of management commentary; (b) simplify or generalize n-grams and lemmatize keywords; (c) expand the list to include inflections and plurals, and alternative English or American spellings; and (d) add disambiguation conditions. Disagreements between coders were reviewed and reconciled. The final list of strategy n-grams comprises 709 elements that nest 2,907 (81%) of the 3,584 words and phrases from the initial list. A full reconciliation from the initial to the final list is available in an online appendix (doi 10.17635/lancaster/researchdata/232).

Appendix B: Using *StratScore* to estimate strategy-related annual report narrative content

This appendix summarizes the mapping from our strategy-related weighted word count metric (*StratScore*) into total report commentary on business model and strategy, and estimates the average incremental increase in strategy-related commentary for Main Market firms (over their AIM counterparts) in response the UK Corporate Governance Code 2010 reporting mandate. We use the following model to project *StratScore* into page content:

$$Pagecount_{i1} = \delta_0 + \delta_1 StratScore_{it} + \omega_{it} \quad (1B)$$

where *Pagecount* is the total number of pages in the narrative component of the annual report relating to fiscal year *t* published by firm *i*, *StratScore* is our report-level weighted frequency of strategy-related words, $\hat{\delta}_1$ is our estimate of the average mapping from strategy words into total report content, $\hat{\delta}_0$ is an estimate of pages in the average annual report whose content is unrelated to business model and strategy, and ω is the regression residual.

Panel A of Table B1 reports findings for equation 1B. The first row presents estimates for the pooled sample using all years. Based on the mean *StratScore* value of 21.3 reported in Panel A of Table 2 for the full sample, the coefficient estimate on *StratScore* of 0.68 for the pooled sample implies that 15 pages of narratives in an average report contain at least some discussion relating to the value creation process. The sample mean number of narrative annual report pages is 28, implying that marginally more than half the pages in the narrative section of the average annual report contain some discussion relating to the value creation process. The final rows of Panel A present results for equation 1B estimated separately using pre-strategy disclosure mandate data for Main Market and AIM firms.

Panel B of Table B1 pairs coefficient estimates on *StratScore* with strategy-related page count for the pre- and post-strategy mandate periods. Findings indicate an implied incremental change in average strategy-related page count for Main Market firms of 11 pages relative to AIM firms (i.e., a 13.98-page increase for Main Market firms compared with a 2.93-page increase for AIM firms).

Table B1: Implied aggregate annual report discussion on the value creation process based on *StratScore*

Panel A: Summary statistics for OLS regressions of narrative page count on *StratScore*

	Intercept	*StratScore*	R^2	N
Pooled sample (all years)	13.699	0.677	0.6349	9,127
	(0.01)	(0.01)		
Main Market sample (pre-mandate)	20.987	0.548	0.5255	3,130
	(0.01)	(0.01)		
AIM sample (pre-mandate)	10.318	0.563	0.3627	3,380
	(0.01)	(0.01)		

Panel B: Implied page count for annual report discussion on the value creation process

		Implied page count for strategy-related content		Δ in implied page count
		CGC2010 = 0	*CGC2010* = 1	
Sample	AIM	5.56	8.49	2.93
	Main Market	14.49	28.47	13.98

Panel A reports coefficient estimates and model summary statistics for narrative page count on *StratScore* (two-tailed probability values reported in parentheses). Panel B uses the *StratScore* coefficient estimates from Panel A in conjunction with the mean *StratScore* value for the corresponding Main Market and AIM samples to estimate the average total page count in the narrative section of the annual report that containing commentary relating to the value creation process. Note the approach does not assume the entire page deals exclusively with strategy- and business model-related issues.

References

Abadie, A., J. Gardeazabal (2003). The economic costs of conflict: A case study of the Basque Country. *American Economic Review* 93(1): 113-132

Acemoglu, D., S. Johnson, A. Kermani, J. Kwak, and T. Mitton (2016). The Value of Connections in Turbulent Times: Evidence from the United States. *Journal of Financial Economics* 121(2): 368-391

ACCA and Eurosif (2013). *What do investors expect from non-financial reporting?* http://www.accaglobal.com/content/dam/acca/global/PDF-technical/sustainability-reporting/tech-tp-wdir.pdf

Accounting Standards Board (2006) *Reporting Statement: Operating and Financial Review* http://www.frc.org.uk/Our-Work/Publications/ASB/UITF-Abstract-24-Accounting-for-start-up-costs/Reporting-Statement-Operating-and-Financial-Review.aspx.

Athanasakou. V., K. Hussainey (2014). The Perceived Credibility of Forward-Looking Performance Disclosures. *Accounting and Business Research* 4: 27–259.

Baginski S., S. Bozzolan, A. Marra, P. Mazzola (2018). Strategic plan presentations to financial analysts and the revision of analysts' earnings forecast, *European Accounting Review* forthcoming

Barney J.B., D.N. Clark. (2007) *Resource-based theory*. *Creating and Sustaining Competitive Advantage*. Oxford University Press

Bilinski, P. D. Lyssimachou, M. Walker (2013) Target price accuracy: International evidence. *The Accounting Review* 88(3): 825-851

Brown, L.D., A. C. Call, M. B. Clement, N. Y. Sharp (2016). The activities of buy-side analysts and the determinants of their stock recommendations. *Journal of Accounting and Economics* 62: 139-156

Department of Business, Skills and Innovation (2010). The Future of Narrative Reporting (https://www.gov.uk/government/consultations/the-future-of-narrative-reporting-a-further-consultation)

Blundell, R., M.C. Dias (2009). Alternative approaches to evaluation in empirical microeconomics. *Journal of Human Resources* 44(3): 565–640

Callen, J.L., M. Khan, H. Lu (2013). Accounting quality, stock price delay, and future stock returns. *Contemporary Accounting Research* 30(1): 269–295.

Casadesus-Masanell, R., J. E. Ricart (2010). From strategy to business models and onto tactics. *Long Range Planning* 43: 195-215

Chen J.V., F., Li, (2015). Discussion of "Textual analysis and international financial reporting: Large sample evidence". *Journal of Accounting and Economics* 50: 181-186

CFA Institute (2006). *Breaking the Short-Term Cycle*. CFA Institute Publication. http://www.cfapubsorg/toc/ccb/2006/2006/1

D'Aveni, R.A., I. C. MacMillan (1990). Crisis and the content of managerial communications: A study of the focus of attention of top managers in surviving and failing firms. *Administrative Science Quarterly* 35(4): 634-657

Darrough, M. N. (1993). Disclosure policy and competition: Cournot vs. Bertrand. *The Accounting Review* 68(3), 534-561

Dahya, J., J. McConnell, N. Travlos (2002). The Cadbury Committee, corporate performance, and top management turnover. *Journal of Finance* 57(1): 461-483

Dechow, P., R. Sloan, A. Sweeney (1996). Causes and consequences of earnings manipulation: an analysis of firms subject to enforcement actions by the SEC. *Contemporary Accounting Research* 13: 1–36

Deloitte (2009). *CFO Insights: The Earnings Guidance Debate*. http://www.deloitte.com

Deloitte (2012). *Joined up writing. Surveying annual reports: This year's story* https://www.iasplus.com/en-gb/publications/uk/other/joined-up-writing-2014-surveying-annual-reports-2012

Diffenbach, J., R. B. Higgins (1987). Strategic credibility can make a difference. *Business Horizons* 30: 13-18

El-Haj, M., P. Alves, P. Rayson, M. Walker, S. Young (2019). Retrieving, classifying and analysing narrative commentary in unstructured (glossy) annual reports published as PDF files. *Accounting and Business Research* forthcoming https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2803275

Ernst and Young (2011). *2011 SEC annual reports – Form 10-K*. Ernst & Young LLP. http://www.ey.com/Publications/vwLUAssets/SECAnnualReports10k_CC0337_November2011/$FILE/SECAnnualReports10k_CC0337_November2011.pdf

European Commission (2017). Commission Guidelines on Non-financial Reporting. https://ec.europa.eu/info/publications/170626-non-financial-reporting-guidelines_en

European Financial Reporting Advisory Group (2013). *The Role of the Business Model in Financial Statements: Research Paper*. European Financial Reporting Advisory Group (EFRAG), the French Autorité des Normes Comptables (ANC) and the UK Financial Reporting Council (FRC). http://www.efrag.org/files/EFRAG%20Output/131218_Business_Model_Research_Paper.pdf

EY (2015). *Annual Reporting in 2014: Reflections on the Past, Direction for the Future.* http://www.ey.com/Publication/vwLUAssets/EY_-_Annual_reporting_in_2014_reflections_on_the_past_direction_for_the_future/$FILE/EY-Annual-reporting-in-2014.pdf

Financial Accounting Standards Board (2001). *Improving Business Reporting: Insights into Enhancing Voluntary Disclosures*. Steering Committee Report, Business Reporting Project. Norwalk, CT.

Financial Reporting Council (FRC). (2009). *Louder than Words: Principles and actions for making corporate reports less complex and more relevant*. Financial Reporting Council: London.

Financial Reporting Council (2010a). *The UK Corporate Governance Code June 2010*. Financial Reporting Council: London. https://www.frc.org.uk/getattachment/31631a7a-bc5c-4e7b-bc3a-972b7f17d5e2/UK-Corp-Gov-Code-June-2010.pdf

Financial Reporting Council (2010b). *Revisions to the UK Corporate Governance Code (Formerly The Combined Code) May 2010*. Financial Reporting Council: London. https://www.frc.org.uk/getattachment/3c363c1c-76ff-4faf-8c7d-7d8aac801cd2/May-2010-report-on-Code-consultation.pdf

Financial Reporting Council (FRC). (2011). *Cutting Clutter in Annual Reports*. Financial Reporting Council: London.

Financial Reporting Council (2012). *Thinking About Disclosures in a Broader Context: A Road Map for a Disclosure Framework*. London: FRC Publications. https://frc.org.uk/Our-Work/Publications/Accounting-and-Reporting-Policy/Thinking-about-disclosures-in-a-broader-contex-File.pdf

Financial Reporting Council (2014). *Guidance on the Strategic Report*. Financial Reporting Council: London https://www.frc.org.uk/getattachment/2168919d-398a-41f1-b493-0749cf6f63e8/Guidance-on-the-Strategic-Report.pdf

Financial Reporting Council (2016). *Lab Project Report: Business Model Reporting*. https://www.frc.org.uk/getattachment/4b73803d-1604-42cc-ab37-968d29f9814c/FRC-Lab-Business-model-reporting-v2.pdf

Grant R.M. (2013). *Contemporary Strategy Analysis*. Wiley (8th Edition)

Gu, F and J.Q. Li (2007). The credibility of voluntary disclosure and insider stock transactions. *Journal of Accounting Research* 45(4): 771-810

Guillaume, M. (2018). Current trends in annual reporting. https://www.reportwatch.net/report-essentials/download/61/

Hail L, C. Leuz (2006). International differences in cost of equity capital: do legal institutions and securities regulations matter? *Journal of Accounting Research* 44: 485-531

Higgins, R.B., J. Diffenbach (1985). The impact of strategic planning on stock prices. *The Journal of Business Strategy* 6: 64-72

House of Commons Treasury Committee (2009). *Banking Crisis: Reforming Corporate Governance and Pay in the City*. The Stationery Office Limited: London. https://publications.parliament.uk/pa/cm200809/cmselect/cmtreasy/519/519.pdf

International Accounting Standards Board (2017). *Disclosure Initiative – Principles of Disclosure*. Discussion Paper PD/2017/1. IFRS Foundation: London

International Accounting Standards Board (2010). *IFRS Practice Statement Management Commentary*. IFRS Foundation. http://www.ifrs.org/Current-Projects/IASB-Projects/Management-Commentary/IFRS-Practice-Statement/Documents/Managementcommentarypracticestatement8December.pdf

International Integrated Reporting Council (2013). *Business Model: Background Paper for <IR>*. International Integrated Reporting Council. http://integratedreporting.org/wp-content/uploads/2013/03/Business_Model.pdf

Kohut, G. F., A. H. Segars (1992). The president's letter to stockholders: An examination of corporate communication strategy. *The Journal of Business Communication* 29(1): 7-21

KPMG (2016). Room for improvement. The KPMG survey of business reporting. KPMG International. https://home.kpmg/xx/en/home/insights/2016/04/kpmg-survey-business-reporting-second-edition.html

Kravet, T. and V. Muslu (2013). Textual risk disclosures and investors' risk perceptions. *Review of Accounting Studies* 18: 1088-1122

Langfiueld-Smith, K. (1997). Management control systems and strategy: A critical review. *Accounting, Organisations and Society* 22(2): 207-232

Lehavy, R., F. Li, K. Merkley (2011). The effect of annual report readability on analyst following and the properties of their earnings forecasts. *The Accounting Review* 86(3): 1087-1115

Lang, M., L. Stice-Lawrence (2015). Textual analysis and international financial reporting: Large sample evidence. *Journal of Accounting and Economics* 60 (2-3): 110-135

Lev, B., F. Gu (2016). *The End of Accounting and the Path Forward for Investors and Managers*. Wiley: New Jersey

Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics* 45: 221–247.

Li, F., R. Lundholm, M. Minnis (2013). A measure of competition based on 10-K fillings. *Journal of Accounting Research* 51(2): 399-436

Lu, H.-Y., J. W. Tucker (2012). Nonearnings corporate guidance. *Financial Management* Winter 947-977

Loughran, T., B. McDonald (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research* 54 (4): 1187-1230

Magretta J. (2012). *Understanding Michael Porter*. Harvard Business Press Books.

Manning C., H. Schütze (1999). *Foundations of Statistical Natural Language Processing*, MIT Press. Cambridge, MA

Palepu, K. G., P. M. Healy, E. Peek (2010). *Business Analysis and Valuation* IFRS edition (2nd edition), South-Western Cengage Learning Inc.

Penman, S.H. (2013). *Financial Statement Analysis and Security Valuation*. Fifth Edition. McGraw Hill. New York

Porter M. (1985). *Competitive Advantage: Creating and Sustaining Superior Performance*. Free Press

PricewaterhouseCoopers (2012). *Trust Through Transparency*. PricewaterhouseCoopers LLP: London

PricewaterhouseCoopers (2007). *The Business Review and AIM*. https://www.pwc.com/gx/en/corporate-reporting/archive/publications/assets/business-review-and-aim.pdf

PwC (2017). *Global Investor Survey on Corporate Reporting*. PwC UK: London. https://www.pwc.com/gx/en/services/audit-assurance/corporate-reporting/investor-view/global-investor-survey-on-corporate-reporting.html

PwC (2016). *Searching for Buried Treasure. A Review of 2015 Strategic Reporting Practices in the FTSE 350*. PwC UK: London.

PwC (2006a). *Guide to forward-looking information. Don't fear the future: communicating with confidence*. PricewaterhouseCoopers LLP: London. https://www.pwc.com/gx/en/audit-services/corporate-reporting/assets/pdfs/860-global-forward-looking-guide.pdf

PwC (2006b). *Operating and Financial Review: Give yourself a head start.* PricewaterhouseCoopers LLP: London. https://www.pwc.com/gx/en/audit-services/corporate-reporting/assets/pdfs/uk-preparers-guide.pdf

Ronda-Pupol, G.A., L.A.N. Guerras-Martin (2012). Dynamics of the evolution of the strategy concept 1962–2008: A co-word analysis. *Strategic Management Journal* 33: 162-188

Rumelt, R. (2011). *Good Strategy Bad Strategy: The Difference and Why it Matters*. Crown Business

Santema, S., J. M. Hoekert, J. Van De Rijt, A. Van Oijen (2005). Strategy disclosure in annual reports across Europe: a study of differences between five countries. *European Business Review* 17(4): 352-366

Santema, S., J. Van De Rijt (2001). Strategy disclosure in Dutch annual reports. *European Management Journal* 19(1): 101-108

Securities and Exchange Commission (2016). *Business and Financial Disclosure Required by Regulation S-K* (Release No. 33-10064; 34-77599; File No. S7-06-16), 13 April 2016, available at: https://www.sec.gov/rules/concept/2016/33-10064.pdf

Securities and Exchange Commission (2013). Report on Review of Disclosure Requirements in Regulation S-K. http://www.sec.gov/news/studies/2013/reg-sk-disclosure-requirements-review.pdf , SEC Offices, Washington D.C.

Sinfield, J.V., E. Calder, B. McConnell, S. Colson (2012). How to identify new business models. *MIT Sloan Management Review* 53(2): 85-90

Stocken, P.C. (2012) Strategic accounting disclosure. *Foundations and Trends in Accounting* 7(4): 197-291

Teufel, S., A. Siddharthan, D. Tidhar (2006). Automatic classification of citation function. In Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing (EMNLP '06). Association for Computational Linguistics, Stroudsburg, PA, USA, 103-110

US Chamber of Commerce (2007). *Commission on the Regulation of US Capital Markets in the 21st Century: Report and Recommendations*

Verrecchia, R.E., 1983. Discretionary disclosure. *Journal of Accounting and Economics* 5, 365–380.

Verrecchia, R. (1980). The rapidity of price adjustments to information *Journal of Accounting and Economics* 2: 63-92

Watts, R., J. Zimmerman (1986). *Positive Accounting Theory*. Prentice Hall: New Jersey

White, M.J. (2013). The path forward on disclosure. Speech given at National Association of Corporate Directors Leadership Conference. http://www.sec.gov/News/Speech/Detail/Speech/1370539878806

Zhang, H., L. Zheng (2011). The valuation impact of reconciling pro forma earnings to GAAP earnings. *Journal of Accounting and Economics* 51(1-2): 186-202

Table 1 Sample selection and descriptive properties

Panel A: Sample selection criteria

		Reports
Population of annual reports between January 2003 and December 2014		15,954
Less:		
Unprocessed annual reports	(1,656)	
Extraction process likely resulting in material error	(3,855)	
Missing identifiers from Datastream	(9)	
Multiple reports during the fiscal year (e.g. change in year ends)	(76)	
Missing accounting and market data for explanatory variables	(1,231)	
Final sample		9,127

Panel B: Distribution of final sample by firms and calendar year

Count of obs. by firm	N	% of total	Count of obs. by year	N	% of total
1	241	13.68	2002	205	2.25
2	231	13.11	2003	609	6.67
3	213	12.09	2004	683	7.48
4	156	8.85	2005	747	8.18
5	150	8.51	2006	857	9.39
6	156	8.85	2007	884	9.69
7	141	8.00	2008	907	9.94
8	153	8.68	2009	847	9.28
9	115	6.53	2010	778	8.52
≥10	206	11.69	2011	790	8.66
			2012	671	7.35
			2013	729	7.99
			2014	420	4.60
Total	1,762	100.00		9,127	100.00

Table 2 Summary statistics for *StratScore* measure strategic commentary and covariates used in large-sample statistical validation tests of *StratScore*. The sample consists of 9,127 firm-year observations for 1,762 firms over the period 2003-2014.

Variables	Mean	Std Dev	Q1	Median	Q3
Panel A: Strategy-related commentary					
Raw strategy count	159	178	45	99	203
Probability weight	0.103	0.007	0.099	0.103	0.107
StratScore	21.278	24.209	5.896	13.023	26.834
StratScore scaled	0.001	0.001	0.001	0.001	0.002
Panel B: Covariates					
CGC2010	0.29	0.45	0.00	0.00	1.00
ABS2006	0.78	0.42	1.00	1.00	1.00
Main	0.46	0.50	0.00	0.00	1.00
Segments	2.24	1.57	1.00	2.00	3.00
Size	10.98	2.32	9.37	10.80	12.47
Financing	0.69	0.46	0.00	1.00	1.00
Competition	-0.81	0.18	-0.98	-0.85	-0.71
ROA	-0.12	0.56	-0.09	0.02	0.07
BM	0.70	1.01	0.26	0.52	0.93
Wordcount_FS	17,066	11,842	9,403	14,440	21,027
Forward	234	203	91	165	312

Panel A reports descriptive statistics for strategy-related annual report content. *Raw strategy count* is the number of strategy words appearing the annual report for firm *i* in year *t*, based on the list of 709 n-grams presented in Appendix A. *Probability weight* is the conditional probability from equation (2) that any given section from the pooled sample of *J* annual reports is as an objectively identified strategy section based on the frequency with which n-gram *k* appears in that section. *StratScore* is our primary measure of strategy-related commentary. It weights counts of the k^{th} n-gram in our list by *Probability weight*. The weighting approach assigns a weight of zero to any n-gram *k* that never appears in the corpus of objectively identified strategy-focused commentary. Weights are also set to zero for any *k* that has a lower proportion of occurrences in the strategy corpus relative to the full annual report narratives corpus. Applying these zero weights reduces from 709 to 231 the number of non-zero weighted elements from *K* and therefore *StratScore* for report *j* equals the weighted sum of frequency counts for 231 n-grams. *StratScore scaled* is equal to *StratScore* divided by total number of words in the financial statements component of the corresponding annual report. Variables reported in Panel B are defined as follows. *CGC2010* is an indicator variable equal to one for annual reports for accounting periods ending on or after 30 June 2011 (post-revision of the UK Corporate Governance Code), and zero otherwise. *ASB2006* is an indicator variable equal to one for annual reports for accounting periods ending on or after 31 March 2006 (post-introduction of the Business Review in annual reports), and zero otherwise. *Main* is an indicator variable equal to one for firms-years registered on the London Stock Exchange Main Market and zero otherwise. *ROA is r*eturn on assets. *Segments* is the number of business segments for firm *i* at the end of year *t*. *Size* is the natural logarithm of total assets in year *t. BM* is the book to market ratio. *Financing* is an indicator variable equal to one if either (1) operating cash flows minus average capital expenditure from years *t* - 3 through to year *t* - 1 (scaled by current assets in *t* - 1) is less than 0.5 (Dechow et al. 1996) or (2) the firm raises capital in t as evidenced by a positive value for proceeds from equity issues (WC04251) or the annual increase in total debt exceeds five percent; and zero otherwise. *Competition* is an industry-year measure equal to minus one multiplied by the sum of the market share (measured by revenue) for industry *j* in fiscal year *t* of the four largest firms ranked by revenue in industry *j*. *WordCount_FS* is the word count for the financial statements component of the annual report, comprising the auditors' report, statement of directors' responsibilities, financial statements and notes to the accounts, statutory shareholder information, statutory five-year summaries, subsidiaries and their locations, and information regarding the annual general meeting (where included). *Forward* is the number of forward-looking words in the narratives component of annual report, all annual sections other than: auditors' report, statement of directors' responsibilities, financial statements and notes to the accounts, statutory shareholder information, statutory five-year summaries, subsidiaries and their locations, and information regarding the annual general meeting (where included).

Table 3. Classification accuracy of *Statscore* relative to gold standard manual classification.

	Domain expert classification		*Stratscore* classification					
Disclosure category	Agree	Disagree	Total classified	Correctly classified	N Type I errors	Precision (%)	N Type II errors	Recall (%)
Low (N = 20)	21		18	16	2	88.89	5	76.19
Medium (N = 20)	10		12	6	6	50.00	4	60.00
High (N = 20)	16		17	14	3	82.35	2	87.50
Total	47	13	47	36	13	73.47	11	76.60

StratScore is our primary measure of strategy-related commentary. It weights counts of the k^{th} n-gram in our list of 709 strategy-related words and phrases by the conditional probability from equation (2) that any given section from the pooled sample of J annual reports is as an objectively identified strategy section based on the frequency with which n-gram k appears in that section. The weighting approach assigns a weight of zero to any n-gram k that never appears in the corpus of objectively identified strategy-focused commentary. Weights are also set to zero for any k that has a lower proportion of occurrences in the strategy corpus relative to the full annual report narratives corpus. Applying these zero weights reduces from 709 to 231 the number of non-zero weighted elements from K and therefore *StratScore* for report j equals the weighted sum of frequency counts for 231 n-grams. "Disclosure category" in column one refers to the relative quality of strategy-related annual report commentary. Twenty reports for each of the three categories are selected at random based on their *Stratscore* ranking for the full sample: reports classified as High are associated with a *Stratscore* value in the top quartile; reports classified as Low are associated with a *Stratscore* value in the bottom quartile; and reports classified as Medium are drawn from all remaining observations. Reports are presented blind (i.e., no information on the reports' original classification and no details of the number of reports selected from each category) to two domain experts who each classify reports into one of the three groups independently. Domain experts' classifications agree for 47 reports (78%), which serve as the gold standard classification against which *Stratscore*'s ability to reproduce similar rankings is assessed. Precision values are equal to the number of true positives divided by the sum of true positives and false positives (multiplied by 100). Recall values are equal to the number of true positives divided by the sum of true positives and false negatives (multiplied by 100).

Table 4 Coefficient estimates and summary statistics for OLS regressions of *StratScore* on factors related to firms' reporting and business environment, and to management incentives. The sample consists of 9,127 firm-year observations for 1,762 firms over the period 2003-2014. Variables are winsorized at the top and bottom 0.5 percentile. Two-tailed *t*-statistics reported in parentheses are based on robust standard errors clustered by firm and year. Superscripts *, ** and *** indicate significance at the 10, five, and one percent levels, respectively.

	Pred.	*StratScore*						*Stratscore scaled*
Variables	sign	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Intercept	?	-66.756***	-49.261***	-43.729***	-56.887***	-12.033**	-4.968	-0.007***
		(-7.39)	(-6.50)	(-6.28)	(-20.64)	(-2.46)	(-1.15)	(-4.25)
CGC2010		12.540**	2.369	3.392	-1.264*	6.616***	6.671***	0.003***
		(2.19)	(1.06)	(1.59)	(-1.65)	(4.88)	(6.69)	(2.94)
Main			1.327	-4.145***	-2.295***	-4.210***	-3.823***	-0.000
			(0.86)	(-3.94)	(-5.70)	(-5.06)	(-5.22)	(-0.01)
CGC2010 x Main	+		15.503***	13.158***	4.942***	5.231***	3.522***	0.004***
			(4.86)	(4.38)	(3.19)	(4.28)	(3.29)	(3.97)
ABS2006	+	3.646***	3.666***	-0.551	-2.996***	1.049*	-0.161	0.000
		(6.55)	(6.47)	(-0.83)	(-4.15)	(1.85)	(-0.34)	(0.36)
ASB2006 x Main				7.981***	3.834***	2.961***	1.917**	0.002*
				(5.19)	(6.85)	(3.23)	(2.18)	(1.90)
Segments	+	0.859***	0.806***	0.822***	0.223*	0.474**	0.570**	0.000*
		(3.60)	(3.27)	(3.35)	(1.77)	(1.98)	(2.48)	(1.89)
Size	+	5.906***	5.019***	5.011***	2.123***	1.428***	0.784***	0.002***
		(10.32)	(12.34)	(12.43)	(10.89)	(5.59)	(4.10)	(15.98)
Financing	+	1.175***	1.731***	1.714***	1.047***	1.059***	0.822**	0.001***
		(2.66)	(3.76)	(3.73)	(3.67)	(2.92)	(2.43)	(3.40)
Competition	–	-4.845	-3.761	-3.585	-3.623**	-6.429***	-5.306**	0.000
		(-1.38)	(-1.28)	(-1.35)	(-2.39)	(-2.70)	(-2.24)	(0.13)
$StratScore_{t-1}$	+				0.691***		-0.053	
					(32.22)		(-0.23)	
ROA		-2.392***	-1.897***	-1.943***	-0.348**	-0.398	-0.626***	-0.000*
		(-5.43)	(-5.07)	(-5.44)	(-2.19)	(-1.40)	(-2.61)	(-1.73)
BM		-2.375***	-1.982***	-1.928***	-0.936***	-0.884***	0.000	-0.001***
		(-4.76)	(-4.67)	(-4.47)	(-3.76)	(-2.94)	(0.68)	(-4.96)
Wordcount_FS		0.000***	0.000***	0.000***	0.000***	0.000**	0.053***	-0.000***
		(4.46)	(4.53)	(4.57)	(2.77)	(2.42)	(17.11)	(-8.04)
Forward						0.073***		
						(19.84)		
LiScore							0.547***	
							(15.28)	
Forward scaled								0.102***
								(5.12)
Time indicators		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry indicators		Yes	Yes	Yes	Yes	Yes	Yes	Yes
N		9,127	9,127	9,127	7,628	9,127	9,217	9,127
Adjusted R^2		0.5207	0.5465	0.5503	0.7380	0.6682	0.6618	0.2800

Table 4 *continued*

The dependent variable in columns 3-7 is the unscaled value of *StratScore* computed using the annual report for fiscal year t published by firm *i*. *StratScore* is equal to the weighted frequency count of 231 n-grams from final strategy wordlist described in Appendix A, where weights are equal to the conditional probability that n-gram *k* is associated with commentaries that unambiguously contain strategy-related content. Results presented in the final column are estimated using *StratScore scaled*, which is equal to *StratScore* divided by total number of words in the financial statements component of the corresponding annual report and multiplied by 10 to aid presentation. Definitions for explanatory variables are as follows. *CGC2010* is an indicator variable equal to one for annual reports for accounting periods ending on or after 30 June 2011 (post-revision of the UK Corporate Governance Code 2010), and zero otherwise. *Main* is an indicator variable equal to one for firms-years registered on the London Stock Exchange Main Market and zero otherwise. *ASB2006* is an indicator variable equal to one for annual reports for accounting periods ending on or after 31 March 2006 (post-introduction of the Business Review in annual reports), and zero otherwise. *Segments* is the number of business segments for firm *i* at the end of year *t*. *Size* is the natural logarithm of total assets in year *t*. *Financing* is an indicator variable equal to one if either (1) operating cash flows minus average capital expenditure from years *t* - 3 through to year *t* - 1 (scaled by current assets in *t* - 1) is less than 0.5 (Dechow et al. 1996) or (2) the firm raises capital in t as evidenced by a positive value for proceeds from equity issues (WC04251) or the annual increase in total debt exceeds five percent; and zero otherwise. *Competition* is an industry-year measure equal to minus one multiplied by the sum of the market share (measured by revenue) for industry *j* in fiscal year *t* of the four largest firms ranked by revenue in industry *j*. $StratScore_{t-1}$ is the lagged value of the dependent variable. *ROA is r*eturn on assets. *BM* is the book to market ratio. *WordCount_FS* is the word count for the financial statements component of the annual report, comprising the auditors' report, statement of directors' responsibilities, financial statements and notes to the accounts, statutory shareholder information, statutory five-year summaries, subsidiaries and their locations, and information regarding the annual general meeting (where included). *Forward* is the number of forward-looking words in the narratives component of annual report, all annual sections other than: auditors' report, statement of directors' responsibilities, financial statements and notes to the accounts, statutory shareholder information, statutory five-year summaries, subsidiaries and their locations, and information regarding the annual general meeting (where included). *LiScore* is the unscaled competition measure proposed by Li et al. (2013) based on frequency counts of “competition”, “competitor”, “competitive”, “compete”, and “competing” including those words with an “s” appended, and then remove any case where “not,” “less,” “few,” or “limited” precedes the word by three or fewer words. *Forward scaled* is equal to *Forward* divided by *WordCount_FS*. All variables are winsorized at the top and bottom 0.5 percentile.

Table 5. Placebo tests regressing non-strategy-related word counts on factors related to firms' reporting and business environment, and to management incentives. The sample consists of 9,127 firm-year observations for 1,762 firms over the period 2003-2014. Two-tailed *t*-statistics reported in parentheses are based on robust standard errors clustered by firm and year. Superscripts *, ** and *** indicate significance at the 10, five, and one percent levels, respectively.

Variable	*Wordcount_FS*	*Wordcount_AR_Resid*
Intercept	-2,501.701	13,712.366***
	(-0.72)	(23.48)
Main	-1,177.777	-360.303***
	(-1.59)	(-3.11)
CGC2010	170.399	75.803
	(0.33)	(0.52)
CGC2010×Main	-1,736.557	500.168
	(-1.49)	(1.44)
ASB2006	-2,602.608***	13.704
	(-6.29)	(0.06)
ASB2006×Main	2,003.069*	425.540
	(1.91)	(1.46)
Segments	571.032***	-20.561
	(4.69)	(-0.43)
Size	2,006.913***	198.687***
	(13.63)	(6.17)
Financing	220.074	196.718**
	(0.80)	(2.28)
Competition	1,045.048	1,177.941**
	(0.78)	(2.45)
ROA	-1,445.725***	-120.128**
	(-5.86)	(-2.45)
BM	-84.868	-15.605
	(-1.53)	(-0.59)
Forward	10.199***	9.758***
	(8.22)	(15.06)
Time indicators	Yes	Yes
Industry indicators	Yes	Yes
Observations	9,127	9,127
Adj. R-squared (%)	46.07	43.87

The dependent variable in column 2 is the word count for the financial statements component of the annual report, comprising the auditors' report, statement of directors' responsibilities, financial statements and notes to the accounts, statutory shareholder information, statutory five-year summaries, subsidiaries and their locations, and information regarding the annual general meeting (where included). The dependent variable in column 3 is the residual word count for the narrative component of the annual report after excluding the following three sections where manual inspection suggests strategy-related commentary is most often located: summary highlights, the letter from the board chair, and management commentary. Definitions for covariates are provided in Tables 3 and 4. All variables are winsorized at the top and bottom 0.5 percentile.

Table 6 Coefficient estimates and summary statistics for OLS regressions of log transformed investor uncertainty proxies on an indicator variable for an exogenous increase in strategy-related commentary. The sample is restricted to include firms with a least two observations in both subperiods. Superscripts *, ** and *** indicate significance at the 10, five, and one percent levels (two-tailed tests), respectively. Two-tailed *t*-statistics (reported in parentheses) are based on robust standard errors clustered by firm and year.

	Investor uncertainty proxy			
Variables	*Forecast dispersion*	*Absolute forecast error (earnings)*	*Absolute forecast error (prices)*	*Bid-ask spread*
Intercept	0.996**	-5.642***	-2.514***	7.114***
	(2.26)	(-15.48)	(-7.78)	(16.11)
Main	0.374***	0.110	-0.167	-0.281
	(3.00)	(0.93)	(-1.54)	(-1.30)
CGC2010	0.077	-0.140	0.079	0.206
	(0.44)	(-0.48)	(0.29)	(0.52)
CGC2010×Main	-0.505***	-0.371**	-0.209	-1.043***
	(-4.49)	(-2.13)	(-1.22)	(-4.11)
ASB2006	-0.241	-0.578**	-0.662**	-0.715**
	(-1.10)	(-2.18)	(-2.19)	(-2.37)
Size	-0.228***	-0.024	-0.146***	-1.056***
	(-5.72)	(-0.72)	(-5.17)	(-31.80)
BM	0.368***	0.125***	0.088**	0.426***
	(7.43)	(2.73)	(2.01)	(4.56)
Segments	0.010	0.029	-0.008	0.032
	(0.39)	(1.05)	(-0.36)	(1.17)
Time	0.019	0.100**	0.045	0.101*
	(0.54)	(2.26)	(0.93)	(1.72)
Wordcount_FS	0.000	-0.000	0.000	0.000
	(0.75)	(-0.44)	(0.68)	(0.41)
Forward	9.275***	4.090	1.135	0.431
	(2.63)	(1.16)	(0.40)	(0.10)
Fog	-0.005	0.005	0.012**	0.011*
	(-0.90)	(0.88)	(2.35)	(1.78)
Industry indicators	Yes	Yes	Yes	Yes
N	4,149	3,947	3,399	6,027
Adjusted R^2	0.1419	0.0650	0.1485	0.7236

The dependent variables in columns 3-6 are defined as follows. *Forecast dispersion* is the standard deviation of individual analyst earnings forecasts from I/B/E/S issued during the financial fiscal year, divided by lagged price. *Absolute forecast error (earnings)* is the absolute value of the forecast error scaled by lagged price, where forecast error is the difference between the I/B/E/S actual EPS and the initial median analyst consensus forecast immediately following the announcement of earnings for previous fiscal year. *Absolute forecast error (prices)* is the absolute value of the forecast error in target prices, where forecast error in target prices is the difference between the target price following announcement of annual earnings and actual stock price 12 months thereafter. *Bid-ask spread* is the average monthly spread (ask minus bid price divided by the average of the bid and ask price) during the financial year divided by lagged price. All four investor uncertainty proxies are log transformed. Definitions for covariates are provided in Tables 3 and 4, with the exception of *Fog* which is the fog index of annual report readability and *Time* which is a linear time-trend variable equal to calendar year minus 2003. The sample is based on 9,127 firm-year observations for 1,762 firms over the period 2003-2014. Sample sizes vary across models according to the data available for the corresponding dependent variable. All variables are winsorized at the top and bottom 0.5 percentile or top 0.5 percentile when bounded at zero.

Table 7 Coefficient estimates and summary statistics for OLS regressions of log transformed investor uncertainty proxies on an indicator variable for increases in *StratScore* post-2010 and a vector of control variables Superscripts *, ** and *** indicate statistical significance at the 10, five, and one percent levels (two-tailed tests), respectively. Two-tailed *t*-statistics (reported in parentheses) are based on robust standard errors clustered by firm and year.

	Investor uncertainty proxy:			
Variables	*Forecast dispersion*	*Absolute forecast error (earnings)*	*Absolute forecast error (prices)*	*Bid-ask spread*
Intercept	-0.619	-3.577**	1.388	-2.767***
	(-0.66)	(-2.52)	(0.88)	(-6.52)
Main	-0.013	-0.370	-0.003	-0.761***
	(-0.05)	(-0.95)	(-0.01)	(-3.90)
CGC2010	-0.060	-0.171	-0.112	-0.139
	(-0.36)	(-0.71)	(-0.50)	(-0.36)
StratScore_increase	-0.317***	-0.396***	-0.305**	-0.735***
	(-2.80)	(-3.09)	(-2.04)	(-3.33)
ASB2006	-0.109	-0.522**	-0.645**	-0.757**
	(-0.45)	(-2.14)	(-2.23)	(-2.55)
Size	-0.332***	0.037	-0.421***	-0.680***
	(-3.15)	(0.41)	(-4.44)	(-7.24)
BM	0.409***	0.098	0.088	0.168**
	(6.77)	(1.29)	(1.26)	(1.98)
Segments	0.065**	0.045	-0.017	0.002
	(2.27)	(1.25)	(-0.52)	(0.10)
Time	-0.010	0.073	0.100**	0.091
	(-0.24)	(1.51)	(2.49)	(1.48)
WordCount_FS	0.000	0.000	-0.000	-0.000
	(0.47)	(0.14)	(-0.82)	(-0.31)
Forward	2.748	-0.385	-0.821	0.547
	(0.98)	(-0.10)	(-0.35)	(0.22)
Fog	0.001	0.005	0.011	0.004
	(0.20)	(1.04)	(1.53)	(0.97)
Industry fixed effects	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes
N	4,149	3,947	3,399	6,027
Adjusted R^2	0.4502	0.3675	0.3154	0.8854

The dependent variables are log transformed proxies for investor uncertainty (see Table 6 for definitions). Explanatory variables are defined in Table 6 with the exception of *StratScore_increase* which is an indicator variable taking the value of one for with the top quartile change in the average value of *StratScore* from the pre- to the post-2010 period. The sample is restricted to firms with a least two observations in both the pre- and post-2010 period. Sample sizes vary across models according to the availability of data for the corresponding dependent variable. All variables are winsorized at the extreme 0.5 percentiles or top 0.5 percentile when bounded at zero.

Table 8 Coefficient estimates and summary statistics for OLS regressions of log transformed stock price delay proxies on an indicator variable for an exogenous increase in strategy-related commentary and an indicator variable for increases in *StratScore* post-2010. Superscripts *, ** and *** indicate statistical significance at the 10, five, and one percent levels (two-tailed tests), respectively. Two-tailed *t*-statistics (reported in parentheses) are based on robust standard errors clustered by firm and year.

Variables	*Stock price delay*	*Stock price delay_fs*	*Stock price delay*	*Stock price delay_fs*
Intercept	6.220***	7.163***	2.781	1.247
	(19.14)	(18.16)	(1.25)	(1.08)
Main	0.035	-0.048	0.137	-0.015
	(0.18)	(-0.30)	(0.36)	(-0.04)
CGC2010	0.504	0.714***	0.582*	0.715***
	(1.21)	(2.58)	(1.82)	(3.45)
CGC2010×Main	-0.478*	-0.319*		
	(-1.76)	(-1.65)		
StratScore_increase			-0.285*	-0.049
			(-1.85)	(-1.27)
ASB2006	-0.923***	-0.901***	-0.878**	-0.904***
	(-2.83)	(-3.06)	(-2.46)	(-3.17)
Size	-0.370***	-0.433***	-0.281***	-0.294***
	(-10.29)	(-12.92)	(-7.92)	(-6.63)
BM	0.168***	0.185***	0.199***	0.166***
	(4.19)	(5.33)	(3.87)	(4.03)
Segments	0.016	0.020	0.009	0.004
	(0.83)	(0.94)	(0.23)	(0.08)
Time	0.168**	0.131***	0.126*	0.116**
	(2.51)	(2.65)	(1.80)	(2.13)
Wordcount_FS	0.000	0.000	-0.000	-0.000
	(1.29)	(0.36)	(-0.08)	(-1.03)
Forward	-7.420***	-6.565**	-7.022***	-4.634
	(-2.90)	(-2.27)	(-2.87)	(-1.42)
Fog	-0.003	-0.004	0.004	0.002
	(-0.38)	(-0.54)	(0.55)	(0.27)
Industry fixed effects	Yes	Yes	Yes	Yes
Firm fixed effects	No	No	Yes	Yes
N	9,061	9,061	6,531	6,531
Adjusted R^2	0.0990	0.1385	0.1577	0.1952

Dependent variables reflect the average delay with which information is impounded into stock prices following the annual report release, calculated using firm-specific regressions of weekly stock returns on contemporaneous weekly market returns and four lags for the 12-month period following publication of the annual report. *Stock price delay* is equal to one minus the ratio of the R^2 from the restricted version of the model (all explanatory variables other than the intercept set to zero) and the R^2 from the corresponding unrestricted model. *Stock price delay_fs* is computed using the same process as used to compute *Stock price delay* but with the four lags of weekly market return replaced by four lags of weekly firm-specific return**.** Both stock price delay variables are log transformed. The sample consists of 9,061 firm-year observations over the period 2003-2014 with available data on stock price delay. Explanatory variables are defined in Table 7. All variables are winsorized at the extreme 0.5 percentiles or top 0.5 percentile when bounded at zero.

Table 9. Coefficient estimates and summary statistics for OLS regressions of log transformed information asymmetry proxies on an indicator variable for post-2010 increases in *StratScore*, conditional low analyst following in the pre-2010 period. Superscripts *, ** and *** indicate statistical significance at the 10, five, and one percent levels (two-tailed tests), respectively. Two-tailed *t*-statistics (reported in parentheses) are based on robust standard errors clustered by firm and year.

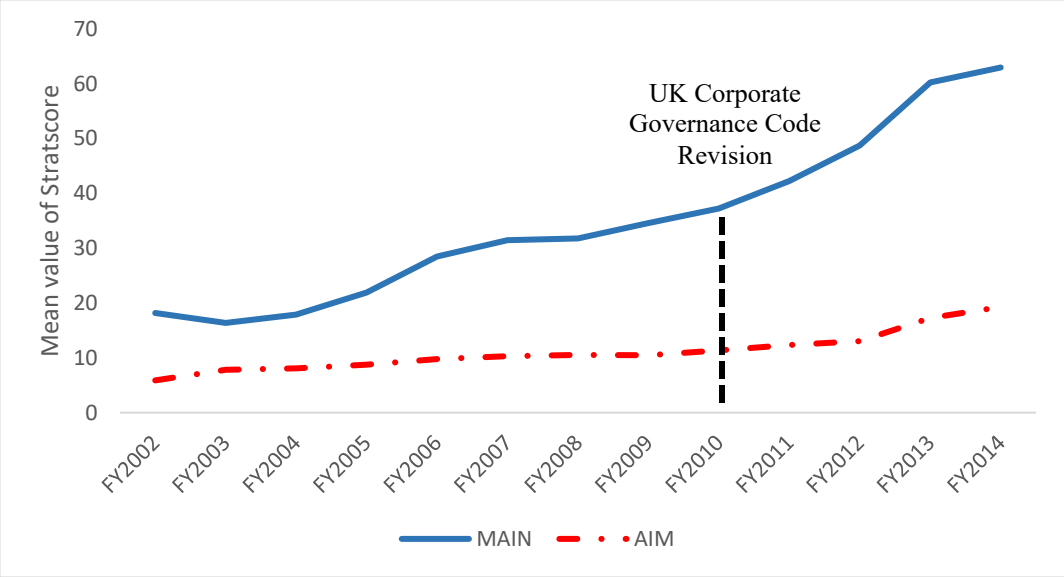
	Absolute forecast error (earnings)	*Absolute forecast error (price)*	*Bid-ask spread*	*Stock price delay*	*Stock price delay_fs*
Intercept	-3.582**	1.388	-0.888	2.762	3.688***
	(-2.52)	(0.88)	(-1.15)	(3.21)	(3.02)
Main	-0.370	-0.003	-0.760***	0.129	-0.023
	(-0.95)	(-0.01)	(-3.89)	(0.34)	(-0.06)
CGC2010	-0.170	-0.112	-0.139	0.581*	0.714***
	(-0.71)	(-0.50)	(-0.36)	(1.81)	(3.42)
StratScore_increase	-0.393***	-0.305**	-0.742***	-0.222	0.012
	(-3.05)	(-2.04)	(-3.33)	(-1.28)	(0.07)
StratScore_increase×LowCover	-1.414***	-1.059***	0.343	-2.256***	-2.186***
	(-8.07)	(-2.67)	(1.10)	(-2.75)	(-3.16)
LowCover	4.137***	0.753*	8.147***	-1.477***	-1.758***
	(9.40)	(1.69)	(16.27)	(-9.44)	(-5.88)
ASB2006	-0.521**	-0.645**	-0.756**	-0.888**	-0.914***
	(-2.14)	(-2.23)	(-2.54)	(-2.47)	(-3.16)
Size	0.037	-0.421***	-0.680***	-0.279***	-0.292***
	(0.41)	(-4.44)	(-7.24)	(-7.91)	(-6.41)
BM	0.099	0.088	0.168**	0.199***	0.165***
	(1.30)	(1.26)	(1.98)	(3.89)	(4.03)
Segments	0.045	-0.017	0.002	0.009	0.004
	(1.25)	(-0.52)	(0.10)	(0.23)	(0.09)
Time	0.073	0.100**	0.091	0.126*	0.116**
	(1.51)	(2.49)	(1.48)	(1.80)	(2.12)
Wordcount_FS	0.000	-0.000	-0.000	-0.000	-0.000
	(0.13)	(-0.82)	(-0.31)	(-0.08)	(-1.02)
Forward	-0.386	-0.821	0.564	-7.115***	-4.724
	(-0.10)	(-0.35)	(0.23)	(-2.86)	(-1.46)

Table 9 *continued*

Fog	0.005	0.011	0.004	0.005	0.003
	(1.04)	(1.53)	(0.96)	(0.57)	(0.29)
Industry Fixed Effects	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes	Yes
N	3,947	3,399	6,027	6,531	6,518
Adj. R-squared	0.3674	0.3154	0.8854	0.1585	0.1960

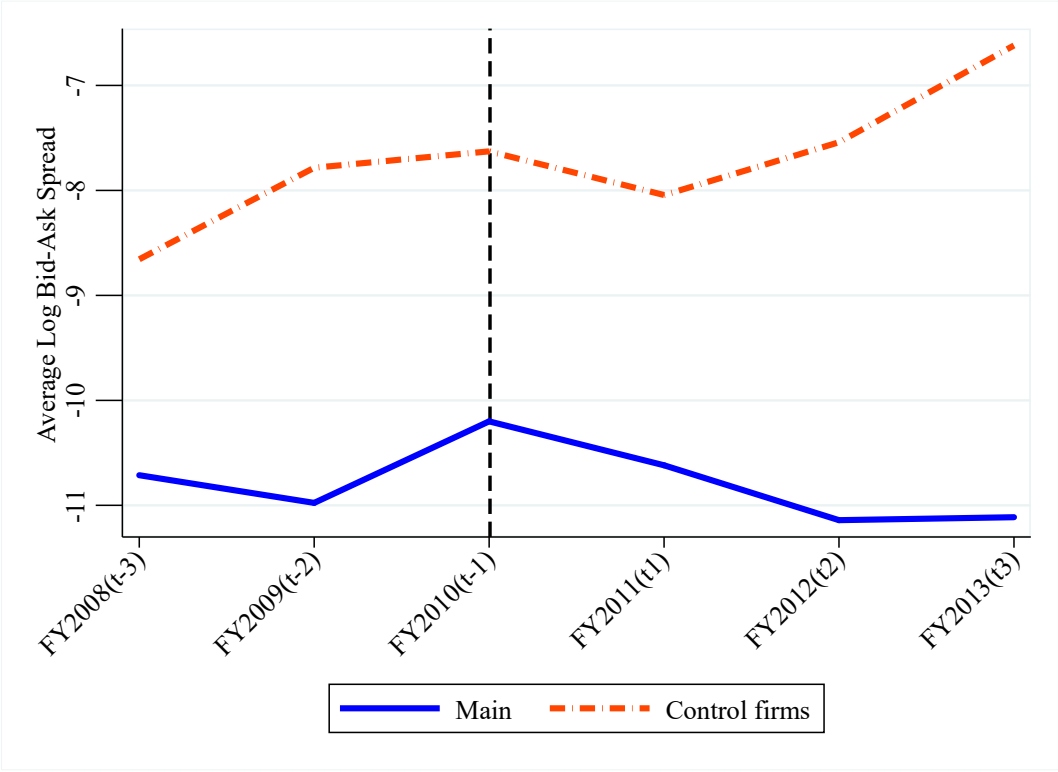
LowCover is an indicator variable taking the value of one for firms whose analyst coverage in the pre-2010 period is in the lowest quartile of analyst following for the sample. Variables are presented in Tables 6 and 8. All variables are winsorized at the extreme 0.5 percentiles or top 0.5 percentile when bounded at zero.

Figure 1 Mean *StratScore* by fiscal year for firms listed on the London Stock Exchange Main Market and Alternative Investment Market (AIM). The sample consists of 9,127 firm-year observations for 1,762 firms over the period 2003-2014.



StratScore is equal to the weighted frequency count of 231 n-grams from final strategy wordlist described in Appendix A, where weights are equal to the conditional probability that n-gram *k* is associated with commentaries that unambiguously contain strategy-related content. The Accounting Standards Board (ASB) issued the best practice statement on management commentary, Reporting Standard 1 *Operating and Financial Review*, encouraging annual report narrative commentary on: corporate objectives and strategy; resources available to deliver those objectives; risks and uncertainties facing the entity; and trends and factors likely to affect the firm's future development (ASB 2006). The best practice guidelines were effective for year-ends ending on or after March 31, 2006 and applied to all LSE-listed firms (Main Market and AIM). Paragraph C.1.2 of the UK Corporate Governance Code (2010), effective for accounting periods beginning on or after June 29, 2010, required LSE Main Market firms (but not their AIM counterparts) to explain in their annual report the basis on which for generating or preserving value over the longer term (the business model), and the strategy for delivering corporate objectives.

Figure 2. Plots of mean bid-ask spread (logged) based on the synthetic control sample method (Abadie and Gardeazabal 2003).



Plots are based on a balanced sample of firms with available data for years surrounding the revision of the UK Corporate Governance Code (2010) effective for accounting periods beginning on or after June 29, 2010 (fiscal year 2011 first effective period of implementation) The solid (blue) line relates to firms listed on the LSE Main market. The dashed (red) line relates to a synthetic control group. Following Acemoglu et al. (2016), each treated unit (i.e. Main market firms in the first year of treatment, i.e. fiscal year 2011) is matched to a combination of untreated units (i.e., AIM firms) that closely match the treated unit over the pre-treatment period (FY2008-FY2010). Matching is performed using a convex weighting matrix that minimizes the Euclidean differences in bid-ask spread on each of the pre-treatment periods. Outcomes for this synthetic control sample, are then projected in the post-treatment period (FY2011-FY2013) using the weights identified by the pre-treatment comparison. These projected outcomes are used as the counterfactual for the treated unit. By construction the differences in bid-ask spreads between the two groups are not significant before FY2011.