Disclosure Similarity and Future Stock Return Comovement

Travis A. Dyer
Cornell University

Darren T. Roulstone
Ohio State University

Andrew Van Buskirk*
Ohio State University

December 2020

Existing research often assumes that firms' financial reporting choices influence their comovement with industry and market returns. We examine the validity of that assumption. We use textual analysis to examine disclosure similarity among firms, and how that similarity influences future stock price comovement. Using three different empirical approaches (samples of firms with amended filings, firms entering/exiting the S&P 500, and firms experiencing changes in analyst coverage due to brokerage closures), we find consistent evidence that firms' disclosure similarity not only predicts, but influences future return comovement. Our results provide empirical support for the idea that firms' disclosure choices influence investors' assessment of return covariances, and thus industry and market betas.

*Corresponding author. We thank Eric Allen, Dirk Black, Jared Flake, Nick Guest, Patrick Kielty, Kyle Peterson, Amy Sheneman, Blake Steenhoven, and workshop participants at the Cornell Accounting Mini-camp, USC Accounting Mini-Conference, and the BYU Accounting Research Symposium for helpful comments. We are grateful to the Samuel Curtis Johnson Graduate School of Management and the Fisher College of Business for financial support.

Disclosure Similarity and Future Stock Return Comovement

Abstract

Existing research often assumes that firms' financial reporting choices influence their comovement with industry and market returns. We examine the validity of that assumption. We use textual analysis to examine disclosure similarity among firms, and how that similarity influences future stock price comovement. Using three different empirical approaches (samples of firms with amended filings, firms entering/exiting the S&P 500, and firms experiencing changes in analyst coverage due to brokerage closures), we find consistent evidence that firms' disclosure similarity not only predicts, but influences future return comovement. Our results provide empirical support for the idea that firms' disclosure choices influence investors' assessment of return covariances, and thus industry and market betas.

1. Introduction

A common assumption in capital market research is that firms' disclosure choices influence their stock return comovement. Examples are abundant. Jin and Myers (2006) directly link opacity and comovement (R^2), noting "An increase in opaqueness, combined with capture by insiders, leads to lower firm-specific risk for investors and higher R^2s." (p. 258). Hutton et al. (2009) show that earnings management, as a measure of opacity, is associated with higher R^2. Brockman et al. (2010) take a similar attitude when they examine the relation between information production and comovement; they conclude that comovement is lower when information production is higher.

These studies and many others are based on the idea that the more firm-specific information the firm produces and disseminates, the less that investors will have to rely upon market-wide information, and the less that the firm's stock price will covary with the market. While this assumption underlies numerous research studies, there is little empirical evidence to support such a causal link – that a firm's discretionary financial reporting choices can influence subsequent return comovement. Our goal is to provide evidence on whether that causal link exists. An affirmative answer would not only validate the assumptions in prior research, but also imply that firms could effectively alter their industry and market betas via their reporting choices, holding their operating decisions and capital structure constant.

In this paper, we study how firms' joint disclosure behavior affects their subsequent return comovement. Specifically, we study the textual similarity of firms' mandated disclosures (10-Ks and 10-Qs) and how that similarity affects future return comovement. Although existing literature documents that similarity in newswire articles and company descriptions *predicts* return comovement (Box 2018; Ibriyamova et al. 2019), we extend that research by assessing whether similarity in firms' disclosures *influences* that comovement.

The obvious challenge in assessing a causal effect of disclosure on returns—particularly over the relatively long windows we are interested in—is that both the firms' disclosures and their returns are influenced by the economic events underlying those disclosures, which would lead to a positive association between disclosure similarity and return comovement even in the absence of a causal relationship. Our study takes three approaches to address this empirical challenge.

First, we focus on a sample of firms that issued both an original and a subsequently-amended filing. For these firms, we calculate two measures of disclosure similarity between amending firms and non-amending firms, one based on the amending firm's original filing (original similarity), and a second based on the firm's amended filing (amended similarity). We consider the amended filing to be more accurate, and therefore the amended similarity to be the "true" similarity measure. We consider the difference between the original similarity and the amended similarity to be the "noise" in the original similarity measure – disclosure similarity that is unwarranted by economic factors.

By construction, the amended filing is not publicly available until after the return comovement is measured, so the firm's amended filing itself cannot influence that comovement. If firms' disclosures influence future return comovement, we would expect the noise in the original similarity measure to predict future return comovement, even though that component of similarity is subsequently shown to be unwarranted. On the other hand, if unobserved economic events drive the observed relation between disclosure similarity and return comovement, we would expect the noise component to have no predictive ability. The more accurate amended similarity would completely explain the return comovement, even though the amended filing was not publicly observable, because the amended filing provides a more accurate description of those economic events.

Our second and third approaches are based on quasi-natural experiments used in prior finance and accounting research to assess causal effects of firm disclosures. Specifically, we examine exogenous changes in investor overlap using S&P 500 inclusions and exclusions, and exogenous changes in analyst coverage based on brokerage closures. In each case, we document that firms respond to the change in investor/analyst composition by altering their disclosure choices; firms' disclosures become more similar to other S&P 500 firms after they enter the S&P 500, and firms' disclosures become less like peer firms' disclosures when brokerage closures lead to a loss of common analyst coverage. We then examine whether these induced changes in disclosure, arguably unrelated to the firms' economic performance, are associated with changes in future return comovement. A positive association would imply that the disclosure changes influenced observed return patterns.

Using SEC filings from 1996 through 2017, we measure disclosure similarity between pairs of firms using the cosine similarity of those firms' 10-Q and 10-K filings. This measure of similarity is driven by a variety of firm-pair characteristics, such as common industry, common analyst coverage, and common owners. Consistent with prior papers showing that similarity in newswire articles and company descriptions predict return comovement (Box 2018; Ibriyamova et al. 2019), disclosure similarity between two firms predicts future stock comovement between those two firms, even after controlling for historical stock comovement, earnings comovement, and a variety of other factors. We extend this prior research by showing that disclosure similarity not only *predicts* but *influences* that future return comovement.

Specifically, our amended filings test shows that the noise component of disclosure similarity from subsequently-amended filings predicts future return comovement, even after controlling for the similarity of the amended filing. The positive predictive ability of *amended*

similarity, despite that amended filing being unobservable during the comovement window, indicates that contemporaneous economic factors subsequently described in the firm's SEC filing explains a significant portion of observed return patterns. However, the incremental predictive ability of similarity *noise*, controlling for those contemporaneous economic factors, implies that the original disclosure itself influenced the subsequent comovement of the firm's stock relative to peer firms.

Our S&P 500 analysis follows a similar approach to Barberis et al. (2005), which documents an increase in return comovement following S&P 500 inclusion and attributes that increase to frictions and investor sentiment. We find that firms' entering the S&P 500 experience an increase in disclosure similarity with other S&P 500 firms; this result suggests that firms alter their disclosure practices in response to a change in analyst/investor constituency (consistent with Jung 2013, Boone and White 2015, and Schoenfeld 2017). We then find that the changes in disclosure similarity around S&P 500 entrance are associated with subsequent changes in return comovement, suggesting that the changes in disclosure influenced the way that investors' perceived the relative value of firms upon addition to the index. In later analysis, we show that the estimated treatment effect on the firm's disclosure similarity is significantly associated with the estimated treatment effect on the firm's return similarity. In other words, when S&P addition induces larger disclosure effects, we observe larger effects on return similarity.

We next incorporate the change in disclosure similarity into the Barberis et al. (2005) research design to assess whether the Barberis et al. (2005) documented increase in return comovement can be attributed to changes in information. We find that the observed change in disclosure similarity does explain a portion of the increased comovement documented by Barberis et al. (2005), again implying that the disclosure changes influenced return patterns. That being

said, the disclosure effect explains only a small portion (roughly 5%) of the effect originally documented by Barberis et al. (2005); thus, our results do not rule out other potential causes offered in Barberis et al. (2005) and subsequent papers.

Our brokerage closure analysis uses the setting studied by Kelly and Ljungqvist (2012) and Balakrishnan et al. (2014), where firms experienced decreases in analyst coverage as a result of brokerage closures; in our setting, we focus on reductions in *shared* analyst coverage. As in the S&P 500 analysis, we find that reductions in shared analyst coverage lead to decreases in disclosure similarity between two firms, suggesting that analysts' preferences influence firms' disclosure choices. We then find that the decreases in disclosure similarity around the brokerage closure are associated with subsequent decreases in return comovement, adding further support to the idea that changes in firms' disclosures lead to changes in observed return comovement.

Overall, our three empirical approaches point to the same conclusion – firms' financial reporting choices influence the extent to which their returns covary with other firms. This conclusion not only validates the assumption made in many prior studies, but also contributes to the fundamental questions of what drives return covariances and how disclosure affects firms' cost of capital. For example, in an influential analytical paper, Lambert et al. (2007) shows how a firm's disclosures can affect its cost of capital via investors' perception of covariances among firms. A necessary link in that disclosure/cost of capital relation is the ability of firm disclosures to causally affect investors' assessed covariances. To the best of our knowledge, ours is the first paper to provide empirical evidence that disclosures do causally affect forward-looking covariances, thus providing support for the Lambert et al. (2007) proposition.

Our paper also contributes to the broad literature studying return comovement. Prior work finds that category investing (Boyer 2011), shared ownership (Antón and Polk 2014), investor

attention (Drake et al. 2016; Huang et al. 2019), and analyst following (Muslu et al. 2014; Israelsen 2016) all play a role in return comovement. We extend this work by showing that disclosure similarity influences future stock return comovement.

Finally, our paper extends the literature on disclosure attributes and their external consequences by showing that a firm's disclosure choices do not just affect that firm's stock price; rather, a firm's *joint* disclosure choices with another firm influence the two firms' subsequent return comovement, even over relatively long periods. In particular, this paper extends the results in Koo et al. (2017), who show that qualitative attributions in firms' earnings forecasts (e.g., attributing an increased earnings forecast to positive industry-trends rather than firm-specific factors) influence the immediate stock price response of peer firms. Our results suggest that firms' joint disclosure choices also influence the way that investors interpret and respond to *future* events, leading to predictable patterns in return comovement.

2. Prior Research and Hypothesis Development

2.1. Return comovement and its interpretation

Return comovement is a fundamental construct in financial economics, underlying beta in the Capital Asset Pricing Model and aiding in portfolio construction more generally. Reflecting this fundamental role, an extensive literature has investigated microeconomic and macroeconomic determinants of return covariance. Central to this research is the idea that return covariances are largely driven by economic forces – firms' stock prices are more likely to move together if those firms are subject to common risk factors (Fama and French 1993) and comovement tends to increase in bad times as firms face similar economic pressure (Erb et al. 1994; Ding et al. 2011).

On top of those common risk factors, prior research documents other determinants of return comovement. Pindyck and Rotemberg (1993) finds greater comovement among stocks predominantly held by institutional investors, suggesting that market segmentation influences comovement. Other papers show that category investing (Boyer 2011), shared ownership (Antón and Polk 2014), investor attention (Drake et al. 2016; Huang et al. 2019), and analyst following (Muslu et al. 2014; Israelsen 2016) all play a role in return comovement. Existing literature also points to behavioral or sentiment-based factors affecting comovement: Barberis et al. (2005) shows that comovement with S&P 500 firms increases after being included in the index, while Green and Hwang (2009) finds that similarly priced stocks comove, suggesting that investors categorize stocks based on index membership and nominal per share prices.

Researchers have interpreted firms' return comovement as being informative about those firms' information environments and disclosure policies. Piotroski and Roulstone (2004) documents that information intermediaries are associated with how much a firm's returns comove with industry and market returns and interprets a lack of market-level synchronicity as an indication that there is more firm-specific information in prices. Jin and Myers (2006) and Brockman et al. (2010) apply a similar interpretation to synchronicity. Haggard et al. (2008) studies analysts' ratings of firms' disclosure policies and finds improved disclosure quality is associated with lower stock price synchronicity (i.e., correlation with market and industry returns), again implying that higher quality disclosure is associated with lower return comovement. Dang et al. (2015) provides country-level evidence that commonality in news coverage is associated with commonality in returns.

Our paper builds on this literature, but extends it in two important ways: First, we examine comovement at the firm-pair level rather than examining how an individual firm's returns covary

with market and industry returns. That is, instead of viewing comovement as the outcome of a single firm's actions or information environment, we evaluate firm-pair comovement as the result of *joint* disclosure characteristics of the firm-pair.

Second, while the existing literature in this area provides ample evidence that firms' disclosure choices, analyst following, and other characteristics are associated with their comovement, an important question is whether these relations are causal. A central idea in Lambert et al. (2007) is that firms' disclosure choices can influence their cost of capital if those disclosures affect how investors perceive the covariance of that firm with other firms. Therefore, knowing whether disclosures causally affect future covariance sheds light on the extent to which firms can influence their own cost of capital through their disclosure choices.[1]

2.2. Disclosure similarity

Just as prior research has studied how firms' stock prices comove, other research investigates how firms' disclosures are similar to one another, and whether those similarities are due to economic forces, managerial discretion, or information intermediaries.[2] Drawing on measures from computational linguistics research (e.g., Salton et al. 1975), Brown and Tucker (2011) studies the similarity of firms' MD&A disclosures from one year to the next. They find that firms experiencing economic shocks are more likely to modify their MD&A language. Hoberg and Phillips (2010) uses 10-K product descriptions to measure product similarity between firms

[1] As evidence that one firm's disclosures can influence investors' perceptions about other firms' cash flows, Lambert et al. (2007) point out that a firm's reported earnings has market and industry components, and that the information transfer literature shows that investors do use one firm's information to revise beliefs about other firms' future earnings and cash flows (pp. 398-399). Our question is not whether one firm's earnings surprise leads to an immediate revision in another firm's stock price, as is typically studied in the information transfer literature. Instead, our question is whether firms' disclosures effect firms' covariances in the subsequent post-disclosure period.

[2] Here, we focus on studies specifically involving disclosure similarity and ignore other aspects of similarity like financial statement similarity (De Franco et al. 2011) and audit similarity (Knechel et al. 2015).

and evaluates how that product similarity influences firms' merger and acquisition decisions. Jung (2013) studies similarity in the context of commodity price risk disclosures finding that firms are more likely to follow a first mover's decision to disclose additional quantitative information when those firms have greater investor overlap.

Three recent papers explore disclosure commonality in the audit context: Drake et al. (2019) shows that auditors engage in disclosure "benchmarking", where they assess nonclient financial statement information, when evaluating a client's financial statement information, resulting in greater comparability among footnotes. De Franco et al. (2018) shows that firms' MD&As are more similar when those firms share the same auditor, and McMullin (2016) provides evidence that financial statement preparers borrow footnote language from firms with whom they share an auditor.

2.3. The relation between disclosure similarity and future return comovement

Two recent papers address the relation between disclosure similarity and future return comovement directly. Box (2018) examines firm-specific newswire content and finds that the linguistic similarity between two firms' content predicts future comovement between those two firms, even after controlling for historical return comovement and other factors. Although the Box (2018) sample is based on newswire stories, much of that content represents firm-issued disclosure. Similarly, Ibriyamova et al. (2019) shows that brief company descriptions from Thomson Reuters and Yahoo! Finance can be used to predict future return comovement. In other words, both studies indicate that qualitative information about two firms is incrementally useful in predicting the future comovement between those firms.

We build on these papers by focusing on similarity in mandatory financial filings, as these filings constitute a significant source of information to investors, are required for all companies with public debt and equity, and contain explicit discussions of forward-looking information (e.g., MD&A). More importantly, while Box (2018) and Ibriyamova et al. (2019) ask whether qualitative textual information can *predict* future return comovement, our question is whether qualitative disclosure similarity, and changes in that similarity, can *affect* future return comovement.

There are several plausible channels through which firms' disclosure choices could affect subsequent return comovement. First, firms could choose to discuss or withhold their exposure to certain risk factors. For example, if a firm's disclosure reveals sensitivity to a particular commodity, the firm's future returns are likely to be more sensitive to news about that commodity, and thus covary more with other firms who have disclosed a sensitivity to that commodity.[3]

Second, the firm could reveal plans to enter into a new line of business, even if that announcement is opportunistic. For example, Cooper et al. (2001) document a substantial "dotcom" effect for firms changing their corporate names to Internet-related dotcom names. More recently, several studies have documented market effects for firms making blockchain-related announcements (e.g., Jain and Jain 2019; Cheng et al. 2019; Cahill et al. 2020; Sharma et al. 2020). In both cases, investors could perceive that those firms are increasingly subject to Internet or blockchain-related factors, leading to greater comovement. Similarly, increasing the amount of disclosure dedicated to industry factors rather than firm-specific factors could lead investors to perceive an increased similarity between the firm and its industry peers, and therefore greater comovement. This latter possibility would be consistent with Koo et al. (2017), who show that

[3] This is true even if those same disclosure choices do not have an immediate pricing effect. Li and Ramesh (2009), for example, find that there is generally little market reaction around quarterly periodic SEC filings. But the Li and Ramesh results do not imply that those filings will be irrelevant for future returns - disclosing a particular risk exposure might not affect the firm's *current* value, but still affect investor response to future news about that risk factor.

peer firms are more likely to experience positive information transfer from earnings forecasts when those earnings forecasts are attributed to industry trends.[4]

Finally, changes in disclosure could influence the nominal categories that investors assign to a particular stock, leading to greater comovement for firms in the same "category". Prior evidence of this phenomenon includes Barberis et al. (2005), who show increased comovement for firms entering the S&P 500; Green and Hwang (2009), who show greater comovement among similarly-priced stocks; Boyer (2011), who shows greater comovement among firms with the same S&P/Barra Value/Growth classification; Chen et al. (2016) who show that otherwise-similar firms have sharply different industry betas depending on whether they fall above or below a 50% sales threshold that determines the firm's SEC-designated industry; and Hameed and Xie (2019), who show that stocks initiating dividends comove with other dividend-paying stocks.

2.4. Research design

Our primary research question is whether two firms' joint disclosure similarity influences their future return comovement. What makes this a difficult question to address is that, even if disclosure similarity has no causal effect on comovement, there will naturally be an association between firms' disclosure similarity and observed return comovement because of the economic forces that drive both disclosure similarity and comovement.

We take three approaches to disentangle a causal effect of disclosure similarity on return comovement. Our first approach examines situations where firms filed an initial SEC filing, and subsequently amended that filing. If the disclosure itself influences future return comovement, we

[4] There are two important distinctions between this paper and Koo et al. (2017). First, Koo et al. (2017) study the immediate short-term response to earnings forecasts, while we are interested in longer-term price behavior following the disclosures in question. Second, we focus on the consequences of two firms' joint disclosure choices, rather than the unidirectional effect of one firm's disclosure on another firm.

expect the disclosure similarity between firm *i*'s initial disclosure and firm *j*'s disclosure to be predictive of future return comovement between firm *i* and firm *j*, even though firm *i*'s initial disclosure was imprecise. On the other hand, if economic factors drive future return comovement and the disclosure itself has no causal effect, we would expect disclosure similarity between firm *i*'s amended (and therefore more accurate) filing and firm *j*'s filing to explain the observed return comovement. If the disclosure similarity based on the original filing—in particular, the unwarranted component of disclosure similarity—predicts future return comovement even after controlling for disclosure similarity based on the amended filing, we would conclude that the disclosure itself influenced future returns.

Our second approach is based on a quasi-exogenous shock to firms' disclosure similarities stemming from their inclusion in (exclusion from) the S&P 500, similar to how Schoenfeld (2017) uses this setting to examine whether voluntary disclosure affects liquidity. Prior literature has shown that firms entering the S&P 500 experience a significant increase in their return comovement with other S&P 500 firms, despite experiencing no obvious change in their economic fundamentals (Barberis et al. 2005). Barberis et al. (2005) argues that this increase in return comovement is driven by common ownership and indicates that frictions or investor sentiment influence comovement in addition to underlying fundamentals.

We exploit the S&P 500 setting based on the assumption that a change in a firm's ownership composition will lead to a change in the firm's disclosure choices, driven by the preferences of different analysts and investors. Prior research supports this assumption. For example, Boone and White (2015) shows that institutional ownership (related to index membership) influences firms' disclosure decisions, leading to a higher level of disclosure. Schoenfeld (2017) shows a similar effect for firms entering the S&P 500, and that this clientele-

driven demand for more information translates into market effects, specifically greater liquidity. Chapman and Green (2018) studies analysts' questions in conference calls and finds that their questions influence managers' future disclosure decisions. Park et al. (2019) shows that common ownership increases disclosure levels because it decreases proprietary costs.

Based on this research, we expect that demand from common owners will cause firms entering the S&P 500 to alter their disclosures so that those disclosures become more similar to other S&P 500 firms, where those same owners were already exerting influence over firm disclosure. If so, we argue that changes in disclosure similarity around S&P 500 inclusion are likely to be driven by changes in their investor composition, rather than changes in those firms' economic fundamentals. Therefore, if these induced increases in disclosure similarity around S&P 500 entrance are associated with increased return comovement, we would conclude that the changes in disclosure similarity influenced that comovement. However, if S&P 500-driven increases in disclosure similarity are not accompanied by an increase in future return comovement, it would suggest that disclosure similarity on its own does not influence return comovement.

Our final approach is based on a similar "disclosure demand" argument. In this case, we exploit the brokerage closures first used by Kelly and Ljungqvist (2012) to study the causal effect of information asymmetry on prices and subsequently used by Balakrishnan et al. (2014) to study the causal effect of firm disclosure on liquidity and cost of capital. The effect of these brokerage closures was to change firms' analyst coverage, and thereby change managers' disclosure behaviors. We expect that if two firms were previously covered by a common analyst, and the brokerage closure eliminates that common analyst coverage, those two firms will subsequently face a lower demand for similar information. As a result, we expect that firms' disclosure similarity will decrease if they experience a loss of common analyst coverage due to the brokerage closures,

but not because of unobservable economic factors. Therefore, if observed decreases in disclosure similarity for affected firms around these closures are associated with decreases in return comovement, we would attribute those changes in comovement to the change in disclosure.

3. Sample Selection and Research Design

3.1. Sample Selection

Our sample begins in 1996, when regulatory filing data for all public firms is first machine-readable via EDGAR, and ends in the third quarter of 2017. We obtain all 10-K and 10-Q filings from EDGAR, making sure to exclude amended filings from the main analysis. Since the filing data comes in a variety of formats (e.g., plain text, HTML), we standardize the documents and make them machine-readable by following the steps in Dyer et al. (2017).[5]

To form our measure of disclosure similarity, we match each firm filing (firm *i*) to a peer firm filing (firm *j*) and constrain the filings to be reporting information for the same calendar quarter. Since the magnitude of possible pairings is extremely large (5,000 firms each quarter over 21 years ≈ 1 billion firm-pairs), we calculate disclosure similarity for a limited number of firm-pairs. To limit the sample, we calculate disclosure similarity for randomly matched firm-pairs. This is done by randomly assigning firms to one of 48 groups and only matching firms within these randomly assigned groups (essentially, we are creating 48 pseudo-industries and randomly assigning firms to one of them). Having matched firm filings within these groups, we estimate the textual similarity of the two filings by using a standard natural language processing technique called cosine-similarity (e.g., Brown and Tucker 2011; Hoberg and Phillips 2010). This process constructs two vectors of words, the first being the words used in the given firm filing and the

[5] See the Appendix for details regarding the 10-K/Q cleaning procedures.

second being the words used in the matched firm filing. Both vectors of words count the frequency of word references in the respective filing. Cosine-similarity takes the cosine of the two vectors of words. The result is a measure of the similarity in word use between the two filings and is our measure of disclosure similarity (*Disclosure Similarity*).

Our sample contains one unique pairing between firm *i* and *j* in each time period *t*. After constraining the sample to have all relevant data from EDGAR, Compustat, CRSP, Thomson Reuters, and I/B/E/S, we obtain a sample of 760,846 unique firm pairs and 9,910,087 observations.

Table 1 presents the descriptive statistics for the sample. For our measures of disclosure similarity and earnings and return comovement, we report descriptives of the raw variables. In our tests, we convert continuous variables to a standard normal for interpretational convenience. The data in Table 1 shows that firm pairs in the sample share a Fama-French 48 industry classification 6% of the time. That is, although we matched firms based on randomly-assigned pseudo industries, the firm-pair shared the same actual industry in about 6% of the observations. Additionally, sharing a common dedicated investor (Bushee 2001), analyst and auditor occurs in 1, 17, and 15% of the sample, respectively.

3.2. Determinants of disclosure similarity

To provide some context for *Disclosure Similarity* for our pseudo-industry sample, we first perform an analysis of its determinants using the following regression specification:

$$Disclosure\ Similarity_{i,j,t} = \beta \cdot Determinants_{i,j,t} + \alpha_{i,t} + \alpha_{j,t} + \varepsilon_{i,j,t} \quad (1)$$

Determinants is a vector of variables we expect to be related to similarity in disclosure, where each variable measures the relation between firm *i* and firm *j* in time *t*. We include fixed effects

associated with firm *i* (firm *j*) in time *t,* which controls for time-specific shocks to disclosure similarity that uniformly affect all the pairings of firm *i* (firm *j*) in time *t*.

We have several predictions regarding the determinants of disclosure similarity, and we broadly classify these determinants into two categories: fundamental and clientele effects. We expect firms who share similar economic fundamentals will use similar language to describe those fundamentals in their periodic filings. For example, we expect firms that share a common industry classification or have a similar reliance on capital markets to disclose similar content. Additionally, extant research suggests that firms who share similar market-to-book, market value of equity, performance, and stock price momentum experience similar return patterns (Fama and French, 1993; Carhart, 1997). As a consequence, we expect the similarities underlying those correlated returns to translate into disclosure similarity. To capture each of these dimensions of fundamental similarity, we identify instances where firm *i* and firm *j* share the same Fama-French 48 industry classification (*Same Industry*). Additionally, we create indicator variables *Similar MTB*, *Similar SIZE*, *Similar ROA*, and *Similar MOM* which take the value of one if firm *i* and firm *j* both share the same decile of market-to-book, market value of equity, return on assets, and stock return, respectively, during the reporting period. We expect each of these determinants to be positively related to *Disclosure Similarity*.

The second group of predictions centers on clientele effects.[6] As discussed earlier, prior research suggests that firm clientele can influence firms' disclosure choices (Boone and White 2015; Schoenfeld 2017; Chapman and Green 2018; De Franco et al. 2018; Park et al. 2019). We study three dimensions of firm clientele: analysts, investors, and auditors. We expect that for each dimension, shared clientele will increase the similarity of two firms' disclosures. *Same Analyst* is

[6] We acknowledge that our two groups of determinants are not mutually exclusive. Industry-focused mutual funds or size-based index funds, for example, will result in some firm pairs having both fundamental and clientele similarity.

an indicator variable taking the value of one if firm *i* and firm *j* both received a forecast from the same analyst during the fiscal period, and zero otherwise. *Same Investor* is an indicator variable taking the value of one if firm *i* and firm *j* both share the same dedicated investor (Bushee, 2001) during the fiscal period, and zero otherwise.[7] *Same Auditor* is an indicator variable taking the value of one if firm *i* and firm *j* both share the same auditor during the fiscal year, and zero otherwise.

Table 2 presents the determinants of disclosure similarity. Column (1) features the variables described above, as well as fixed effects for each firm *i* quarter (Firm *i* x Yearqtr FE) and each firm *j* quarter (Firm *j* x Yearqtr FE). These fixed effects capture each firm's average similarity across all other firms they are matched with, in each quarter. The first key takeaway from this table is that dimensions of fundamental similarity are related to disclosure similarity. Specifically, sharing the same industry is associated with a 0.244 standard deviation increase in disclosure similarity, the largest effect in the table. Secondly, measures of similarity in common risks are also positively associated with disclosure similarity – two firms are more likely to share common disclosure language if they are in the same decile of *Size*, market-to-book (*MTB*), profitability (*ROA*), and momentum (*MOM*).

The second takeaway is that there seem to be economically and statistically significant clientele effects. Sharing an analyst or a dedicated institutional investor with a peer firm is associated with higher disclosure similarity. These associations broadly support the notion that common analysts and investors encourage commonality in disclosure. That being said, the associations do not necessarily imply causality; disclosure similarity (or fundamental similarity underlying that disclosure similarity) is likely to attract common analyst and investor coverage, which would lead to the same association without clientele influencing disclosure. We also find

[7] We focus on dedicated investors, rather than transient and quasi-indexers, since we expect dedicated investors to be more likely to influence the disclosure of the firm.

that sharing an auditor is positively related to disclosure similarity, as documented by De Franco et al. (2018).

We next perform the same regression, this time including additional fixed effects for each firm *i*,*j* pair (Firm *i* x Firm *j*), and thus capturing intertemporal variation within each firm pair. Column (2) shows the results of this regression. In comparison to Column (1), the results here show a substantial decline in both the magnitude and the statistical significance of the independent variables, with each estimated coefficient decreasing by about 90% from Column (1) to Column (2). The general message from these two columns is that the determinants of disclosure similarity are largely based on the firm-pair, and do not vary much within that firm-pair over time.

4. Empirical Results – disclosure similarity and future return comovement

4.1. Main Sample Analysis

We next evaluate the association between *Disclosure Similarity* and future return comovement within our broad sample, where *Future Return Comovement*$_{i,j,t}$ is the Pearson correlation of the stock returns of firm *i* and *j* over the 90 days following the later date of firm *i* and firm *j*'s disclosure. We start by graphically showing the univariate relation between return comovement and three intuitive determinants. Specifically, Figure 1 shows *Future Return Comovement* by quartile of disclosure similarity, earnings comovement, and lagged return comovement. Unsurprisingly, past return comovement shows the sharpest relation with future return comovement. Earnings comovement shows a positive, but small, association with future return comovement. Finally, disclosure similarity shows a clear positive association with future return comovement.

We next measure the determinants of return comovement in a multivariate setting using the following OLS regression:

$$Future\ Return\ Comovement_{i,j,t} = \beta_1 \cdot Disclosure\ Similarity_{i,j,t} + Controls \qquad (2)$$

Disclosure Similarity is as previously defined and represents the degree of qualitative similarity between two firms. Higher degrees of correlation (β_1) suggest that *Disclosure Similarity* can be used to incrementally predict *Future Return Comovement* between two firms. *Controls* represents a vector of control variables that includes previous return comovement during the fiscal quarter (*Return Comovement*$_{i,j,t}$) and quarterly earnings comovement over the prior three years (*Earnings Comovement*$_{i,j,t}$). In addition to these pair-wise control variables, we impose a variety of fixed effects to evaluate the robustness and pervasiveness of these results.

Table 3 presents the results of this multivariate regression. Column (1) controls for historical stock return comovement, historical earnings comovement, and both firm *i*'s and firm *j*'s average disclosure similarity during that year. Column (2) includes an additional set of firm-pair fixed effects, which leaves only intertemporal variation within firm-pairs to be explained. As expected, in Column (1) we see that historical return comovement (*Return Comovement*$_{i,j,t}$) is an economically significant predictor of future return comovement, with a correlation of 0.135 (i.e., a one-standard deviation increase in past return comovement is associated with a 0.135 standard deviation increase in future return comovement).[8] Notably, this relation is no longer significant in Column (2), which implies that return comovement for a given firm pair is fairly sticky over time, and thus is absorbed by the firm-pair fixed effects. Both cross-sectionally and intertemporally (Columns (1) and (2)), we find that *Disclosure Similarity* is incrementally predictive of subsequent return comovement, controlling for the previous periods return comovement as well as earnings

[8] This auto-correlation in return comovement becomes larger when fixed effects are omitted. Disclosure similarity continues to be a significant predictor of future return comovement with fixed effects omitted.

comovement. Columns (3) and (4) show that, regardless of the fixed effects structure, *Disclosure Similarity* is positively associated with *Future Return Comovement* when we control for commonality in firm characteristics.

The results in Table 3 are broadly in line with the results in Box (2018) and Ibriyamova et al. (2019), who show that similarity in newswire text and company descriptions can be used to predict future return comovement. While this analysis demonstrates that *Disclosure Similarity* can be useful in forming more efficient portfolios, the documented association could be caused by two very different factors. First, the disclosure similarity metric might simply reflect shared underlying fundamentals for the two companies that are not captured by the other control variables. In this case, disclosure does not necessarily have a causal effect on future return patterns. Alternatively, the common disclosure may cause investors to adjust their perceptions of how similar the companies are (as in Lambert et al. 2007), and therefore lead to more similar returns. We are interested in this second possibility – whether disclosure similarity *causally* affects future return comovement.

4.2. Causal Inference -- Amended Filing Test

Our first causal inference setting focuses on amended SEC filings, where we evaluate firm *i*-firm *j* disclosure similarity based on both firm *i*'s original filing and firm *i*'s amended filing. We assume that firm *i*'s original filing captures the firm's fundamentals with some amount of noise, and that the amended filing better reflects firm *i*'s economic fundamentals. Therefore, disclosure similarity based on the original filing (*Original Disclosure Similarity*) should be a noisier measure of the economic similarity between firm *i* and firm *j* than disclosure similarity based on the amended filing (*Amended Disclosure Similarity*).

If unobservable economic fundamentals rather than firm disclosures are driving our earlier results, we would expect that *Amended Disclosure Similarity* better predicts future return comovement and subsumes any relation between *Original Disclosure Similarity* and future return comovement. On the other hand, if *Original Disclosure Similarity* predicts future return comovement even after controlling for *Amended Disclosure Similarity*, we would conclude that the (noisier) disclosure itself is influencing future return comovement, rather than only reflecting the underlying fundamentals. In particular, we focus on the component of *Original Disclosure Similarity* that was unwarranted by economic fundamentals and refer to that component as *Disclosure Similarity – Error*, equal to *Original Disclosure Similarity* minus *Amended Disclosure Similarity*.

To test these predictions, we estimate the following specification:

$$\begin{aligned} Future\ Return\ Comovement_{i,j,t} = \ & \alpha_1 \cdot Amended\ Disclosure\ Similarity_{i,j,t} \\ & + \alpha_2 \cdot Disclosure\ Similarity - Error_{i,j,t} \\ & + Controls \end{aligned} \tag{3}$$

A positive and significant α_2 estimate would indicate that return comovement is associated with the noise component of disclosure similarity and would imply that disclosure similarity influences comovement in future returns.

We obtain a sample of amended 10-K/Q filings that occur 90 days after the initial filing date. This design choice means that the amended filings themselves could not have influenced the observed return comovement, as they are filed after return comovement is measured. We constrain amended filings to be longer than the original disclosure to remove cases where the amended disclosure is not likely to be a complete 10-K/Q disclosure. The resulting sample of original and amended filings is paired with peer firms' filings that occurred during the same time period.

Table 4 presents the association between future return comovement and disclosure similarity using the amended filing sample. Both columns present the association between *Future Return Comovement* and the two measures of disclosure similarity (*Disclosure Similarity – Error* and *Disclosure Similarity – Amended)*, with Column (2) featuring a richer set of control variables. We find that both components of disclosure similarity (i.e., the warranted similarity reflected in the amended measure and the unwarranted noise component) are related to future return comovement patterns, with the estimated coefficient for the noise component equal to roughly 60% of the estimated coefficient for the warranted component. These results provide further evidence that disclosure similarity does not simply reflect economic links that are reflected in observed return patterns independent of the disclosure. Instead, we view these results as evidence that *Disclosure Similarity* between two firms causally affects the trading pattern between those two firms.[9]

4.3. Causal inference – S&P 500 entrance/exit

As discussed earlier, prior research on return comovement documents that clienteles influence return comovement patterns. In this section, we explore the role of disclosure similarity in these settings. We do this in two steps. First, we evaluate whether clientele causally influences disclosure similarity. Second, we evaluate the extent to which shifts in disclosure similarity are a channel through which index membership and common analyst coverage increase future return comovement.

[9] An alternative design choice is to simply include both the Original similarity measure and the Amended similarity measure, rather than focusing on the "error" portion of the original disclosure. Untabulated results indicate that the two approaches yield the same inferences.

We start by identifying 511 firms that either entered or exited the S&P 500 between 1996 and 2017 and also had requisite data for control variables (firm *i*). We then calculate the firm's *Disclosure Similarity* with a peer firm (firm *j*) in each period *t,* where peer firms are a combination of (1) every firm in the S&P 500 at that time and (2) 1,000 randomly selected firms.[10] The benefit to benchmarking against a set of randomly selected firms is that it allows us to rule out the concern that the new S&P 500 firms were changing their disclosures in such a way as to become more similar to *all* firms around the time they were entering the index (e.g., adding more common or boilerplate language to their disclosures). *Treat* is an indicator variable taking the value of one if firm *i* is paired with an S&P 500 firm. *Inclusion* is an indicator if firm *i* is in the S&P 500 index in period *t*. The interaction of *Treat* and *Inclusion* tells us if the firm's disclosure patterns tend to converge (diverge) with their S&P 500 peers upon entering (exiting) the S&P 500 index. We impose firm-time effects for both firm *i* and firm *j*. These fixed effects subsume the main effects of *Post* and *Treat* in this setting.

Table 5 Panel A presents the results of this analysis. We find a positive and significant coefficient on *Treat x Inclusion*, indicating that upon entry into (exit from) the S&P 500 index, treated firms start disclosing more (less) similarly to their index peers. In Column (2), we augment the analysis by controlling for similarity in quantitative firm fundamentals as well as other clientele effects. These controls are designed to control for convergence in quantitative fundamentals that might be driving the change in *Disclosure Similarity* around inclusion into (exclusion from) the S&P 500. With this specification, we continue to find a positive and significant coefficient on *Treat x Inclusion*, consistent with *Disclosure Similarity* increasing significantly upon inclusion

[10] We require the set of firms available for random selection to be firms that survive more than half the sample period. This constraint increases the number of matches with data in each period and allows for time-series comparisons. We find similar results when drawing from the entire population of firms and randomly selecting new firms each period.

into the index. This result suggests that firm disclosures tend to converge in response to sharing a common investor clientele. Given that fundamentals are not likely to have changed dramatically upon inclusion into the S&P 500, this result also highlights the possibility that increased disclosure similarity may contribute to the increased return comovement documented by Barberis et al. (2005) upon entry to the S&P 500. (We discuss the relation between our results and the Barberis et al. (2005) results in Section 4.5.)

We argue that the results in Table 5 Panel A establish that entrance to or exit from the S&P 500 induces changes in disclosure similarity that are unrelated to changes in firms' economic fundamentals. We next assess whether these induced changes in disclosure similarity (i.e., the change in similarity immediately around the firm's entrance/exit from the index) are associated with changes in future return comovement using the following regression:

$$\Delta Future\ Return\ Comovement_{i,j,t} = \beta_1 \cdot \Delta Disclosure\ Similarity_{i,j,t} + Controls \qquad (4)$$

The results of this regression are shown in Table 5 Panel B. Column (1) shows the results of the regression when we exclude the change in common investor, while Column (2) includes the change in common investor. We present both specifications because the change in common investor is likely to have two effects, only one of which we are interested in. First, an increase in common ownership can directly increase return comovement due to non-disclosure reasons; if an institutional investor owns large stakes in two companies, significant inflows and outflows to that fund can lead to common price pressure on those companies' stocks, increasing their comovement (e.g., Greenwood and Thesmar, 2011). Second, an increase in common ownership can increase common disclosure, which then leads to increased comovement. Our interest is in the second effect, and we expect that regressing the change in return comovement on the observed change in disclosure similarity should capture this indirect effect. Nonetheless, we include the change in

common investor variable in Column (2) to ensure that there are no correlated omitted factors that may be influencing our results.

In both cases, we observe a positive and significant coefficient on the change in disclosure similarity ($p<0.01$). We interpret these estimates as evidence that disclosure similarity has a causal effect on future return comovement – as two firms' disclosures feature increasingly common language, investors perceive their values as being increasingly correlated.

It is important to note that, even though the sample focuses on the periods immediately surrounding the firms' S&P 500 entrance/exit, the firm might naturally experience changes in disclosure similarity that are unrelated to that entrance/exit. If so, we could still observe a positive association between changes in disclosure similarity and changes in return comovement unrelated to the index entrance/exit; our earlier results indicate that disclosure similarity is correlated with future return comovement outside of this natural experiment. Although we cannot conclusively rule out this possibility, we can compare the observed relation in our natural experiment (Columns 1 and 2) to the same observed relation in the general population (Column 3). If the change in disclosure caused by S&P 500 inclusion is unrelated to fundamentals *and* disclosure has no causal effect on return comovement, we expect that the relation between disclosure similarity and comovement (i.e., the estimated β_1 coefficient) will be significantly smaller than in the full population. That is, an irrelevant increase in *Disclosure Similarity* will have the same effect on the observed coefficient as adding random noise to any independent variable – the estimated coefficient will be biased downward.

In this case, though, we observe the opposite – the relation in our natural experiment (0.0195) is approximately 22 times higher than the relation in the general population (0.0009). We interpret that stark difference as evidence that the changes in disclosure similarity driven by the

S&P 500 inclusion/exit are largely substantive (i.e., changing the content of the information disseminated) and thus likely to influence investors' perceptions. In contrast, we expect that changes in disclosure similarity in the broader population are often uninformative or unlikely to influence investors' perceptions, such as when firms' disclosures change because they have changed auditors and adopted different language styles for their footnotes. Therefore, we interpret the comparative results as evidence that changes in disclosure similarity driven by membership in the S&P 500 have a causal effect on firms' future return comovement.

4.4. Causal inference – Brokerage Closures

Our third approach to causal inference is similar in structure to the S&P 500-related analysis. We first establish that a natural experiment—brokerage closures in this case—influenced firms' disclosure similarities for arguably non-fundamental reasons. We then assess whether the change in disclosure similarity around those events is correlated with the change in future return comovement for affected firms. Finally, we compare the magnitude of the correlation around the brokerage closures to the magnitude of the correlation for the general population.

We identify firm pairs (firms *i*, *j*) that were followed by a common analyst in one of the closed brokerage firms identified by Kelly and Ljungqvist (2012). The sample includes 18,536 firm pairs that lost common analyst coverage as a result of these closures, which we label *Treat* pairs.[11] As a benchmark to these treated pairs, we include all firm *i* and firm *j* pairs from the main sample, which represent our non-treated pairs. *Post* is an indicator variable if the analyst brokerage firm has closed. If firms respond to common analyst coverage with increases in disclosure

[11] Specifically, we focus on firm-pairs that lost their only common analyst as a result of the brokerage closure.

similarity, then we would expect a negative response to the interaction of *Treat* and *Post*, when the firm pair has lost that common coverage.

Table 6 Panel A presents the results of this difference in difference test. We find that the loss of common analyst coverage is associated with decreased disclosure similarity in firm pairs. The magnitude of this effect is equivalent to a 0.076 standard deviation decrease in disclosure similarity upon losing a common analyst. Column (2) adds controls for changes in fundamentals and other clientele effects. Under this more stringent specification, we continue to find that disclosure similarity decreases after losing a common analyst. This evidence suggests that firm disclosures respond to common analyst coverage, and the loss thereof.

Table 6 Panel B shows the results of the regression shown in equation (4), where the change in future return comovement is regressed on the change in disclosure similarity around the brokerage closures. As in the S&P 500 analysis, we present our results both excluding and including the direct effect of common analyst coverage. Column (1) excludes the common analyst coverage variable (*ΔSame Analyst*) and shows a positive and significant ($p<0.01$) association between the change in disclosure similarity and the change in future return comovement; the observed decreases in disclosure similarity around brokerage closures are associated with decreases in future return comovement. Column (2) includes the common analyst coverage variable and shows a similar result, with comparable magnitudes.

As in the prior section, we then compare the magnitude of the estimated correlation in the natural experiment to the magnitude of the correlation in the general population (Column 3). As before, we find that the size of the association in our experiment is substantially larger than in the general population (roughly ten times the size). We interpret this relative effect as evidence that

the brokerage closures induced substantive decreases in disclosure similarity that led to a decrease in investors' perceived covariances between the two firms.

4.5. Disclosure similarity as a channel for previously documented return comovement

We next examine the changes in disclosure similarity and changes in return comovement around the two natural experiments using a different empirical structure, with the goal of identifying a causal channel for previously-documented changes in return comovement. Specifically, we first estimate the change in return comovement around each natural experiment without accounting for the contemporaneous change in disclosure similarity. We then estimate the same regression, this time including the contemporaneous change in disclosure similarity. If the disclosure similarity effect subsumes part of the treatment effect, we interpret that result as evidence that disclosure similarity is a causal channel through which these natural experiments increased return comovement. We take this approach for two reasons. First, it strengthens our claim that disclosure similarity has a causal effect on return comovement. Second, it allows us to benchmark the size of that effect relative to the more direct effects of these natural experiments on future return comovement.

We show the results of these analyses in Table 7, with the S&P 500 entrance/exit in Panel A and the brokerage closures in Panel B. Column (1) of Panel A shows a significant increase in future return comovement upon entrance to the S&P 500 index, in line with the results documented in Barberis et al. (2005) using S&P 500 changes from earlier periods. Column (2) shows the effect of including *Disclosure Similarity*, which is also significantly associated with future return comovement. Including Disclosure Similarity leads to a decrease in the estimated entrance/exit effect (*Treat x Inclusion*) by about 0.005, or 4.87% of the main effect in Column (1). We interpret

these results as evidence that part of the S&P 500 effect documented in Barberis et al. (2005) is information based, driven by changes in disclosure similarity. However, the disclosure effect is fairly small relative to the more direct effect that changes in investor composition have on future return comovement.

Panel B shows a similar analysis around brokerage closures. Column (1) shows that those brokerage closures induced a decrease in return comovement among the affected firms. Column (2) shows that *Disclosure Similarity* contributes to this effect ($p<0.01$). The decrease in the treatment effect from Column (1) to Column (2) is -0.003, or about 3.66% of the main effect in Column (1). Our inference here is similar to Panel A: The change in disclosure similarity contributed to a decline in return comovement around brokerage closures, but the size of that effect pales in comparison to the more direct effect of losing common analyst coverage.

Across both the changes models described in Sections 4.3 and 4.4 and the seemingly unrelated regression specifications described in this section, we find evidence consistent with disclosure similarity playing an indirect role in the comovement effects documented in prior studies.

4.6. Additional analysis – estimating the intensity of S&P 500 effect

Our analysis of S&P 500 changes described in Table 5 relied upon the simple change in observed disclosure around S&P entrance/exit to explain observed changes in subsequent return comovement. To more precisely estimate the causal effect of index-induced disclosure on return comovement, we follow a similar approach to that in Schoenfeld (2017) who uses index changes to study the effect of disclosure on stock liquidity. Specifically, we first estimate the observed changes in both disclosure similarity and future return comovement as a function of firm and time

period fixed effects. We then remove those estimated fixed effects from the variable of interest (disclosure similarity or future return comovement) and estimate the residual variable as a function of the index entrance/exit treatment effect. These results are shown in Table 8, Panel A, where the treatment effect (*Treat x Inclusion*) has a positive and significant association with both *Disclosure Similarity* (Column 1) and *Future Return Comovement* (Column 2).

In Panel B, we evaluate the relation between future return comovement and disclosure similarity using a difference-in-differences changes model, similar to Schoenfeld (2017). This analysis is analogous to an intensity of treatment effect, where we compare cross-sectional variation in the estimated treatment effect on disclosure similarity (*TxI Disclosure Similarity*) with that of the estimated treatment effect on future return comovement (*TxI Future Return Comovement*). In both Column (1) and Column (2), we find a significantly positive association between the intensity of change in disclosure similarity and the intensity of the change in subsequent return comovement. These results further corroborate our conclusion that disclosure similarity causally influences return comovement.

5. Conclusion

We use textual analysis to examine disclosure similarity among firms, and how that similarity not only predicts, but influences future stock price comovement. Across three different empirical settings, we find evidence that disclosure similarity (and changes in disclosure similarity) between two firms causally affects the future comovement between those two firms' stock prices. Moreover, we show that this disclosure effect explains a portion (albeit a minority) of the previously documented change in return comovement for firms entering the S&P 500 (Barberis et al. 2005).

To the best of our knowledge, our study provides the first empirical evidence that firms' disclosures causally affect future return comovement. As a result, we provide validation for a common assumption in prior research (e.g., Jin and Myers 2006; Hutton et al. 2009; Brockman et al. 2010) and provide novel empirical support for the proposition in Lambert et al. (2007) that firms' disclosure choices can influence their cost of capital by affecting investors' assessed covariances. Our paper contributes to the broad literature studying return comovement, and shows that disclosure similarity (like common ownership, analysts, and auditors) both predicts and influences future stock return comovement. Finally, our paper extends the literature on disclosure attributes and their consequences by showing that a firm's disclosure choices do not just affect that firm's stock price; rather, a firm's *joint* disclosure choices with another firm influence the two firms' subsequent return comovement.

References

Antón, M., Polk, C., 2014. Connected Stocks. *Journal of Finance* 69 (3): 1099–1127.

Balakrishnan, K., Billings, M. B., Kelly, B., Ljungqvist, A., 2014. Shaping Liquidity: On the Causal Effects of Voluntary Disclosure. *Journal of Finance* 69 (5): 2237–2278.

Barberis, N., Shleifer, A., Wurgler, J., 2005. Comovement. *Journal of Financial Economics* 75 (2): 283–317.

Boone, A. L., White, J. T., 2015. The Effect of Institutional Ownership on Firm Transparency and Information Production. *Journal of Financial Economics* 117 (3): 508–533.

Box, T., 2018. Qualitative Similarity and Stock Price Comovement. *Journal of Banking & Finance* 91: 49–69.

Boyer, B. H., 2011. Style-related comovement: Fundamentals or labels? *The Journal of Finance* 66 (1): 307-332.

Brockman, P., Liebenberg, I., Schutte, M., 2010. Comovement, Information Production, and the Business Cycle. *Journal of Financial Economics* 97 (1): 107–129.

Brown, S. V., Tucker, J. W., 2011. Large-Sample Evidence on Firms' Year-over-Year MD&A Modifications. *Journal of Accounting Research* 49 (2): 309–346.

Bushee, B.J., 2001. Do institutional investors prefer near-term earnings over long-run value? *Contemporary Accounting Research* 18 (2): 207-246.

Bushee, B. J., Noe, C. F., 2000. Corporate Disclosure Practices, Institutional Investors, and Stock Return Volatility. *Journal of Accounting Research* 38 (3): 171–202.

Cahill, D., Baur, D. G., Liu, Z., Yang, J. W., 2020. I Am a Blockchain Too: How Does the Market Respond to Companies' Interest in Blockchain? *Journal of Banking & Finance* 113: 105740.

Carhart, M. M., 1997. On persistence in mutual fund performance. *Journal of Finance* 52 (1): 57-82.

Chapman, K., Green, J. R., 2018. Analysts' influence on managers' guidance. *The Accounting Review* 93 (1): 45-69.

Chen, H., Cohen L., Lou, D. 2016. Industry Window Dressing. *Review of Financial Studies* 29 (12): 3354–3393.

Cheng, S. F., De Franco, G., Jiang, H., Lin, P., 2019. Riding the Blockchain Mania: Public Firms' Speculative 8-K Disclosures. *Management Science* 65 (12): 5901–5913.

Dang, T.L., Moshirian, F., Zhang, B., 2015. Commonality in News around the World. *Journal of Financial Economics* 116: 82-110.

De Franco, G., Fogel-Yaari, H., Li, H., 2018. MD&A Textual Similarity and Auditors. *Working Paper*.

De Franco, G., Kothari, S. P., Verdi, R. S., 2011. The Benefits of Financial Statement Comparability. *Journal of Accounting Research* 49 (4): 895–931.

Ding, R., Zhou, H., Li, Y., 2019. Social Media, Financial Reporting Opacity, and Return Comovement: Evidence from Seeking Alpha. *Journal of Financial Markets, forthcoming*.

Drake, M. S., Jennings, J., Roulstone, D. T., and Thornock, J. R., 2016. The Comovement of Investor Attention. *Management Science* 63 (9): 2847–2867.

Drake, M. S., Lamoreaux, P. T., Quinn, P. J., and Thornock, J. R., 2019. Auditor Benchmarking of Client Disclosures. *Review of Accounting Studies, Forthcoming*.

Dyer, T., Lang, M., and Stice-Lawrence, L., 2017. The Evolution of 10-K Textual Disclosure: Evidence from Latent Dirichlet Allocation. *Journal of Accounting & Economics* 64 (2/3): 221–245.

Erb, C. B., Harvey, C. R., Viskanta, T. E., 1994. Forecasting International Equity Correlations. *Financial Analysts Journal* 50 (6): 32–45.

Fama, E. F., French, K. R., 1993. Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics* 33 (1): 3–56.

Fama, E. F., French, K. R., 1997. Industry Costs of Equity. *Journal of Financial Economics* 43 (2): 153–193.

Green, T. C., Hwang, B., 2009. Price-based return comovement. *Journal of Financial Economics* 93 (1): 37-50.

Haggard, K. S., Martin, X., Pereira, R., 2008. Does voluntary disclosure improve stock price informativeness? *Financial Management* 37 (4): 747-768.

Hoberg, G., Phillips, G. M., 2010. Product Market Synergies and Competition in Mergers and Acquisitions: A Text-Based Analysis. *Review of Financial Studies* 23 (10): 3773–3811.

Huang, S., Huang, Y., Lin, T., 2019. Attention allocation and return co-movement: Evidence from repeated natural experiments. *Journal of Financial Economics* 132 (2): 369-383.

Hutton, A. P., Marcus, A. J., Tehranian, H., 2009. Opaque financial reports, R2, and crash risk. *Journal of Financial Economics* 94 (1): 67-86.

Ibriyamova, F., Kogan, S., Salganik-Shoshan, G., Stolin, D., 2019. Predicting Stock Return Correlations with Brief Company Descriptions. *Applied Economics* 51 (1): 88–102.

Israelsen, R.D, 2016. Does common analyst coverage explain excess comovement? *Journal of Financial and Quantitative Analysis* 51 (4): 1193-1229.

Jain, A., Jain, C., 2019. Blockchain Hysteria: Adding "Blockchain" to Company's Name. *Economics Letters* 181: 178–181.

Jin, L., Myers, S. C., 2006. R2 around the world: New theory and new tests. *Journal of Financial Economics* 79 (2): 257-292.

Jung, M. J., 2013. Investor Similarity and Diffusion of Disclosure Practices. *Review of Accounting Studies* 18 (1): 167–206.

Kelly, B., Ljungqvist, A., 2012. Testing Asymmetric-Information Asset Pricing Models. *Review of Financial Studies* 25 (5): 1366–1413.

Knechel, W. R., Vanstraelen, A., Zerni, M., 2015. Does the Identity of Engagement Partners Matter? An Analysis of Audit Partner Reporting Decisions. *Contemporary Accounting Research* 32 (4): 1443–1478.

Koo, D. S., J. Julie Wu, and P. E. Yeung. 2017. Earnings Attribution and Information Transfers. *Contemporary Accounting Research* 34 (3): 1547–1579.

Lambert, R., Leuz, C., Verrecchia, R. E., 2007. Accounting information, disclosure, and the cost of capital. *Journal of Accounting Research* 45 (2): 385-420.

McMullin, J. L., 2016. Can I Borrow Your Footnotes? Footnote Boilerplate's Learning Externality. *SSRN eLibrary*.

Muslu, V., Rebello, M., Xu, Y., 2014. Sell-Side Analyst Research and Stock Comovement. *Journal of Accounting Research* 52 (4): 911–954.

Park, J., Sani, J., Shroff, N., White, H., 2019 Disclosure incentives when competing firms have common ownership. *Journal of Accounting and Economics* 67 (2-3): 387-415.

Pindyck, R. S., Rotemberg, J. J., 1993. The Comovement of Stock Prices. *Quarterly Journal of Economics* 108 (4): 1073–1104.

Piotroski, J. D., Roulstone, D. T., 2004. The Influence of Analysts, Institutional Investors, and Insiders on the Incorporation of Market, Industry, and Firm-Specific Information into Stock Prices. *The Accounting Review* 79 (4): 1119–1151.

Salton, G., Wong, A., Yang, C. S., 1975. A Vector Space Model for Automatic Indexing. *Communications of the ACM* 18 (11): 613–620.

Schoenfeld, J., 2017. The Effect of Voluntary Disclosure on Stock Liquidity: New Evidence from Index Funds. *Journal of Accounting and Economics* 63 (1): 51–74.

Sharma, P., Paul, S., Sharma, S., 2020. What's in a Name? A Lot If It Has "Blockchain." *Economics Letters* 186: 108818.

Appendix

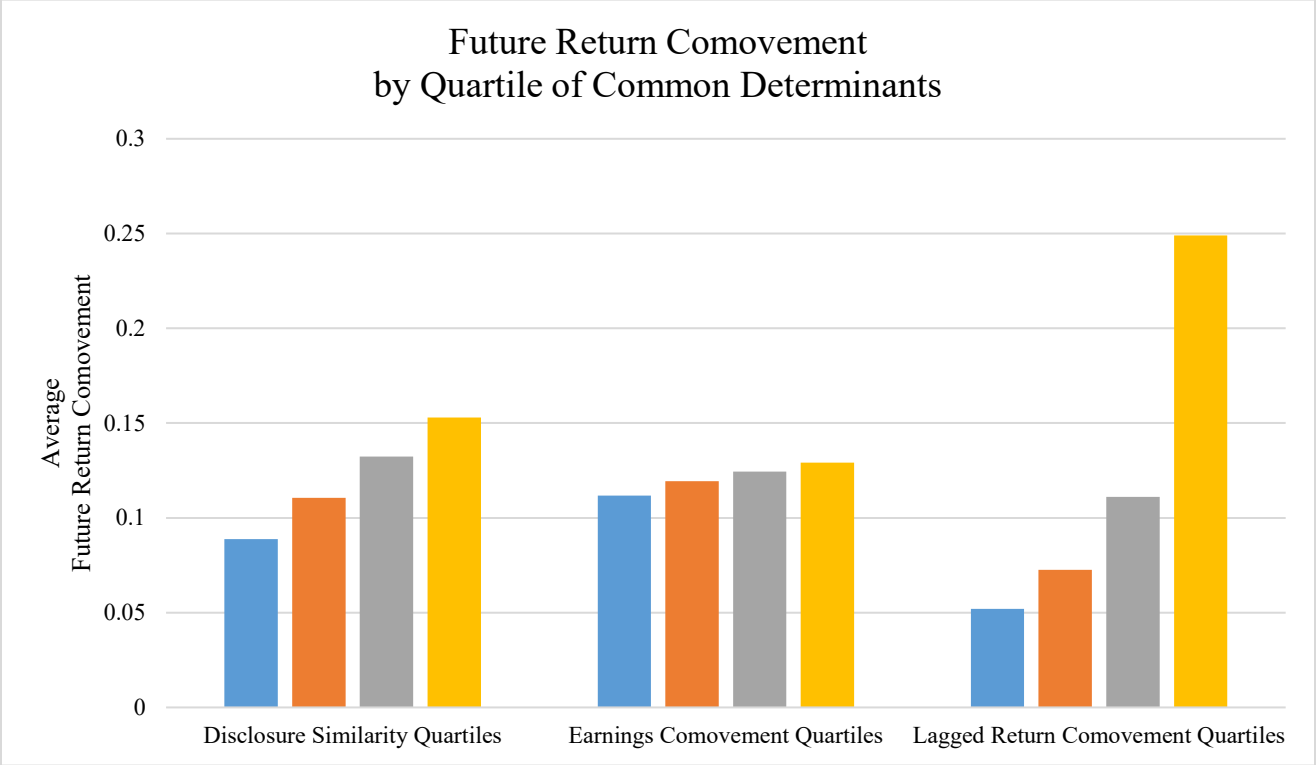
Variable Definitions:

Variable Name	Variable Definitions
Disclosure Similarity	The cosine-similarity of a firm's 10-K/Q report with that of a peer firm's 10-K/Q report, where peer reports come from the same calendar quarter.
Future Return Comovement	The Pearson correlation coefficient between a firm's stock returns (firm *i*) and a peer firm's stock return (firm *j*) over the 90 days following the later of firm *i* and firm *j*'s disclosure.
Return Comovement	The Pearson correlation coefficient between a firm's stock returns (firm *i*) and a peer firm's stock return (firm *j*) over the calendar quarter prior to 10-K/Q disclosure.
Earnings Comovement	The Pearson correlation coefficient between a firm's quarterly earnings and a peer firm's quarterly earnings over the three years prior to 10-K/Q disclosure.
Same Industry	An indicator variable taking the value of one if firm *i* and firm *j* both share the same Fama-French 48 industry classification during the fiscal quarter, and zero otherwise (Fama and French, 1997).
Same Investor	An indicator variable taking the value of one if firm *i* and firm *j* both share the same dedicated investor (Bushee, 2001) during the fiscal period, and zero otherwise.
Same Analyst	An indicator variable taking the value of one if firm *i* and firm *j* both received a forecast from the same analyst during the fiscal period, and zero otherwise.
Same Auditor	An indicator variable taking the value of one if firm *i* and firm *j* both share the same auditor during the fiscal year, and zero otherwise.
Similar MTB	An indicator variable taking the value of one if firm *i* and firm *j* both share the same decile of Market-to-Book, and zero otherwise. Market-to-Book is measured as the firm Market Value of Equity scaled by Book Value of Equity as of the end of the fiscal quarter.
Similar SIZE	An indicator variable taking the value of one if firm *i* and firm *j* both share the same decile of Market Value of Equity, and zero otherwise. Market Value of Equity is measured as of the end of the fiscal quarter.
Similar ROA	An indicator variable taking the value of one if firm *i* and firm *j* both share the same decile of Return on Assets, and zero otherwise. Return on Assets is measured as net income scaled by total assets as of the end of the fiscal quarter.
Similar MOM	An indicator variable taking the value of one if firm *i* and firm *j* both share the same decile of Stock Return, and zero otherwise. Stock Return is measured as the percentage change in stock price over the fiscal quarter.

10-K/Q Cleaning Procedure

We follow the steps from Dyer et al. (2017) in cleaning 10-K/Q documents. First, we remove all header and appendix information. This content includes the SEC header section at the start of all 10-K documents, as well as any graphics, zip files, XML files, excel files, 101 exhibits, 100 exhibits, pdf files, and XBRL. Next, we remove HTML tagging from the text using the HTML::Parser Perl module. We further remove remaining tags (e.g., <TEXT>, <PAGE>, <DOCUMENT>, <TYPE>) and lines with certain tags (e.g., <S>, <C>) following Miller (2010). Next, we exclude lines with fewer than 20 characters or 15 alphanumeric characters. Lastly, we exclude paragraphs with (1) more than 50% non-alphabetic characters and (2) fewer than 80 characters.

Figure 1



Notes: This figure presents the average value of future return comovement by quartile of disclosure similarity, earnings comovement, and the preceding return comovement. Each observation in the data represents a pairing of firms (firms *i* and *j*) during a calendar quarter (time *t*). Firms are randomly assigned to fixed groupings across time and are only paired to firms within this grouping in order to reduce the magnitude of possible pairings. *Future Return Comovement* is measured as the Pearson correlation coefficient between a firm's stock returns (firm *i*) and a peer firm's stock return (firm *j*) over the 90 days following the later of firm *i* and firm *j*'s disclosure. *Disclosure Similarity* is measured as the cosine-similarity of a firm's 10-K/Q report with that of a peer firm's 10-K/Q report, where peer reports come from the same calendar quarter. *Earnings Comovement* is the Pearson correlation coefficient between a firm's quarterly earnings and a peer firm's quarterly earnings over the three years prior to 10-K/Q disclosure. *Return Comovement* is the Pearson correlation coefficient between a firm's stock return (firm *i*) and a peer firm's stock return (firm *j*) over the calendar quarter prior to 10-K/Q disclosure.

Table 1 - Descriptive Statistics

	N	Mean	Std Dev	Q1	Median	Q3
*Disclosure Similarity**	9,889,293	0.93	0.04	0.91	0.94	0.95
*Future Return Comovement**	9,889,293	0.12	0.19	-0.01	0.10	0.23
*Return Comovement**	9,889,293	0.12	0.18	-0.01	0.10	0.23
*Earnings Comovement**	9,889,293	0.03	0.39	-0.25	0.02	0.30
Same Industry	9,889,293	0.06	0.23	0	0	0
Same Investor	9,889,293	0.01	0.08	0	0	0
Same Analyst	9,889,293	0.17	0.38	0	0	0
Same Auditor	9,889,293	0.15	0.36	0	0	0
Similar MTB	9,889,293	0.11	0.31	0	0	0
Similar SIZE	9,889,293	0.11	0.32	0	0	0
Similar ROA	9,889,293	0.11	0.31	0	0	0
Similar MOM	9,889,293	0.10	0.30	0	0	0

Notes: This table presents the descriptive statistics for our main sample. Each observation in the data represents a pairing of firms (firms *i* and *j*) during a calendar quarter (time *t*). Firms are randomly assigned to fixed groupings across time and are only paired to firms within this grouping in order to reduce the magnitude of possible pairings. *Disclosure Similarity* is measured as the cosine-similarity of a firm's 10-K/Q report with that of a peer firm's 10-K/Q report, where peer reports come from the same calendar quarter. *Future Return Comovement* is measured as the Pearson correlation coefficient between a firm's stock returns (firm *i*) and a peer firm's stock return (firm *j*) over the 90 days following the later of firm *i* and firm *j*'s disclosure. *Return Comovement* is the Pearson correlation coefficient between a firm's stock return (firm *i*) and a peer firm's stock return (firm *j*) over the calendar quarter prior to 10-K/Q disclosure. *Earnings Comovement* is the Pearson correlation coefficient between a firm's quarterly earnings and a peer firm's quarterly earnings over the three years prior to 10-K/Q disclosure. *Same Industry* is an indicator variable if firm *i* and firm *j* share the same Fama-French 48 classification. *Same Analyst* is an indicator variable taking the value of one if firm *i* and firm *j* both received a forecast from the same analyst during the fiscal period, and zero otherwise. *Same Investor* is an indicator variable taking the value of one if firm *i* and firm *j* both share the same dedicated investor (Bushee, 2001) during the fiscal period, and zero otherwise. *Same Auditor* is an indicator if firm *i* and firm *j* share the same auditor during the period. *Similar MTB*, *Similar SIZE*, *Similar ROA*, *Similar MOM* are indicator variables taking the value of one if firm *i* and firm *j* share the same decile of Market-to-Book, Market Value of Equity, Return on Assets, and Stock Return, respectively during the prior fiscal quarter, and zero otherwise. *Continuous variables are not converted into standard normal in this table.

Table 2 - Determinants of Disclosure Similarity

	Disclosure Similarity					
	(1)			(2)		
	Coef.	*t-Stat*		*Coef.*	*t-Stat*	
Same Industry	0.244	22.82	***	0.016	3.04	***
Same Investor	0.046	20.34	***	0.005	3.80	***
Same Analyst	0.141	18.19	***	0.016	3.96	***
Same Auditor	0.027	13.97	***	0.002	1.30	
Similar MTB	0.023	12.31	***	0.002	2.73	***
Similar SIZE	0.024	15.38	***	0.003	4.47	***
Similar ROA	0.053	22.06	***	0.003	5.15	***
Similar MOM	0.015	9.17	***	0.001	1.18	
Firm *i* x Yearqtr FE	Yes			Yes		
Firm *j* x Yearqtr FE	Yes			Yes		
Firm *i* x Firm *j*	No			Yes		
R-squared	0.75			0.82		
Obs	9,889,293			9,820,666		

Notes: This table presents the correlation between disclosure similarity and measures of firm commonality. Each observation in the data represents a pairing of firms (firms *i* and *j*) during a calendar quarter (time *t*). Firms are randomly assigned to fixed groupings across time and are only paired to firms within this grouping in order to reduce the magnitude of possible pairings. *Disclosure Similarity* is measured as the cosine-similarity of a firm's 10-K/Q report with that of a peer firm's 10-K/Q report, where peer reports come from the same calendar quarter. It is further standardized by subtracting off the mean and dividing by the standard deviation. *Same Industry* is an indicator variable if firm *i* and firm *j* share the same Fama-French 48 classification. *Same Analyst* is an indicator variable taking the value of one if firm *i* and firm *j* both received a forecast from the same analyst during the fiscal period, and zero otherwise. *Same Investor* is an indicator variable taking the value of one if firm *i* and firm *j* both share the same dedicated investor (Bushee, 2001) during the fiscal period, and zero otherwise. *Same Auditor* is an indicator if firm *i* and firm *j* share the same auditor during the period. *Similar MTB, Similar SIZE, Similar ROA, Similar MOM* are indicator variables taking the value of one if firm *i* and firm *j* share the same decile of Market-to-Book, Market Value of Equity, Return on Assets, and Stock Return, respectively during the prior fiscal quarter, and zero otherwise. Singleton observations are dropped. Standard errors are clustered by firm *i* and year-quarter.

Table 3 - Future Return Comovement and Disclosure Similarity

	Future Return Comovement											
	(1)			(2)			(3)			(4)		
	Coef.	*t-Stat*		*Coef.*	*t-Stat*		*Coef.*	*t-Stat*		*Coef.*	*t-Stat*	
Disclosure Similarity	0.034	15.39	***	0.002	2.93	***	0.020	15.14	***	0.002	2.65	***
Return Comovement	0.135	13.58	***	0.008	1.44		0.119	12.98	***	0.007	1.29	
Earnings Comovement	0.009	17.96	***	0.003	6.05	***	0.006	14.16	***	0.002	5.43	***
Same Industry							0.107	13.97	***	0.032	5.22	***
Same Investor							0.206	11.71	***	0.069	7.91	***
Same Analyst							0.384	25.86	***	0.082	11.70	***
Same Auditor							0.037	9.38	***	0.008	3.62	***
Similar MTB							0.027	15.30	***	0.014	12.67	***
Similar SIZE							0.145	24.18	***	0.052	19.71	***
Similar ROA							0.030	17.16	***	0.010	10.20	***
Similar MOM							0.023	12.00	***	0.011	8.34	***
Firm *i* x Yearqtr FE	Yes			Yes			Yes			Yes		
Firm *j* x Yearqtr FE	Yes			Yes			Yes			Yes		
Firm *i* x Firm *j* FE	No			Yes			No			Yes		
R-squared	0.54			0.59			0.55			0.59		
Obs	9,889,293			9,820,666			9,889,293			9,820,666		

Notes: This table presents the predictive ability of disclosure similarity in forecasting future return comovement. Each observation represents a pairing of firms (firms *i* and *j*) during a calendar quarter (time *t*). Firms are randomly assigned to fixed groupings across time and are only paired to firms within this grouping in order to reduce the magnitude of possible pairings. *Future Return Comovement* is measured as the Pearson correlation coefficient between a firm's stock return (firm *i*) and a peer firm's stock return (firm *j*) over the 90 days following the later of firm *i* and firm *j*'s disclosure. *Disclosure Similarity* is measured as the cosine-similarity of a firm's 10-K/Q report with that of a peer firm's 10-K/Q report, where peer reports come from the same calendar quarter. *Return Comovement* is the Pearson correlation coefficient between a firm's stock returns (firm *i*) and a peer firm's stock return (firm *j*) over the calendar quarter prior to 10-K/Q disclosure. *Earnings Comovement* is the Pearson correlation coefficient between a firm's quarterly earnings and a peer firm's quarterly earnings over the three years prior to 10-K/Q disclosure. All continuous variables are standardized by subtracting off the mean and dividing by the standard deviation. All other variables are as previously defined. Singleton observations are dropped. Standard errors are clustered by firm *i* and year-quarter.

Table 4 - Future Return Comovement and Disclosure Similarity: Amended Filings

	Future Return Comovement					
	(1)			(2)		
	Coef.	*t-Stat*		*Coef.*	*t-Stat*	
Disclosure Similarity - Amended	0.035	14.78	***	0.026	13.03	***
Disclosure Similarity - Error	0.022	8.07	***	0.017	6.79	***
Return Comovement	0.104	11.30	***	0.098	11.25	***
Earnings Comovement	0.008	8.31	***	0.006	7.34	***
Same Industry				0.023	7.09	***
Same Investor				0.147	18.59	***
Same Analyst				0.025	10.68	***
Same Auditor				0.021	8.69	***
Similar MTB				0.090	11.47	***
Similar SIZE				0.058	9.73	***
Similar ROA				0.345	7.22	***
Similar MOM				0.048	7.75	***
Firm *i* x Yearqtr FE	Yes			Yes		
Firm *j* x Yearqtr FE	Yes			Yes		
R-squared	0.45			0.45		
Obs	10,516,065			10,516,065		

Notes: This table presents the association between disclosure similarity and future return comovement. Each observation represents a pairing of firms (firms *i* and *j*) during a calendar quarter (time *t*). In this sample, each firm *i* amended their original filing more than 90 days after the original release. *Future Return Comovement* is measured as the Pearson correlation coefficient between a firm's stock returns (firm *i*) and a peer firm's stock return (firm *j*) over the 90 days following the later of firm *i* and firm *j*'s disclosure. *Disclosure Similarity* (*Disclosure Similarity - Amended*) is measured as the cosine-similarity of a firm's original (amended) 10-K/Q report with that of a peer firm's original 10-K/Q report, where peer reports come from the same calendar quarter. *Disclosure Similarity - Error* is the difference between *Disclosure Similarity* and *Disclosure Similarity - Amended*. All continuous variables are standardized by subtracting off the mean and dividing by the standard deviation. All other variables are as previously defined. Standard errors are clustered by firm *i* and year-quarter.

Table 5 - Relation between Future Return Comovement and Disclosure Similarity around S&P 500 Inclusion/Exclusion

Panel A - Changes in Disclosure Similarity around S&P 500 Inclusion/Exclusion

	Disclosure Similarity					
	(1)			(2)		
	Coef.	*t-Stat*		*Coef.*	*t-Stat*	
Treat x Inclusion	0.015	3.98	***	0.013	3.62	***
Same Industry				0.098	14.41	***
Same Analyst				0.105	17.22	***
Same Auditor				0.009	4.13	***
Similar MTB				0.009	6.14	***
Similar SIZE				0.008	6.51	***
Similar ROA				0.018	10.66	***
Similar MOM				0.004	1.76	*
Firm *i* x Yearqtr FE	Yes			Yes		
Firm *j* x Yearqtr FE	Yes			Yes		
R-Squared	0.83			0.83		
Observations	6,724,766			6,724,766		

Table 5 - Continued

Panel B – Changes in Future Return Comovement and Disclosure Similarity around S&P 500 Inclusion/Exclusion

	Δ Future Return Comovement								
	S&P 500 Inclusion/Exclusion			*S&P 500 Inclusion/Exclusion*			*Main Sample*		
	(1)			(2)			(3)		
	Coef.	*t-Stat*		*Coef.*	*t-Stat*		*Coef.*	*t-Stat*	
Δ Disclosure Similarity	0.020	14.11	***	0.020	13.89	***	0.001	2.26	**
Δ Earnings Comovement	0.014	4.51	***	0.014	4.33	***	-0.003	-3.86	***
Δ Same Industry	0.053	1.30		0.052	1.27		0.018	1.24	
Δ Same Investor				0.088	15.44	***	0.103	62.02	***
Δ Same Analyst	-0.021	-1.39		-0.021	-1.44		0.006	0.99	
Δ Same Auditor	0.026	1.26		0.026	1.25		0.016	4.07	***
Δ Similar MTB	0.009	2.14	**	0.009	2.06	**	0.001	0.70	
Δ Similar SIZE	0.029	4.45	***	0.029	4.46	***	0.004	3.33	***
Δ Similar ROA	0.001	0.32		0.001	0.33		0.001	1.12	
Δ Similar MOM	0.010	2.79	***	0.010	2.80	***	-0.005	-5.49	***
R-Squared	0.0007			0.0007			0.0005		
Observations	386,501			386,501			8,296,799		

Notes: Panel A evaluates how disclosure similarity changes upon a firm's inclusion into (exclusion from) the S&P 500. For each firm that entered or exited the S&P 500 during our sample period (firm *i*), we compare their disclosure with other S&P 500 firms and randomly selected non S&P 500 firms (firm *j*) during that time period. *Treat* is an indicator variable taking the value of one if firm *i* is paired with an S&P 500 firm *j* and zero otherwise. *Inclusion* is an indicator if firm *i* is in the S&P 500 in period *t*. All other variables are as previously defined. All continuous variables are standardized by subtracting off the mean and dividing by the standard deviation. Standard errors are clustered by firm *i* and year-quarter. Panel B evaluates the relation between future return comovement and disclosure similarity using a changes model. Columns (1) and (2) evaluate the relationship between future return comovement and disclosure similarity in the two quarters after being included in (excluded from) the S&P 500 index. Column (3) presents benchmark results using the main sample. All other variables are as previously defined. All continuous variables are standardized by subtracting off the mean and dividing by the standard deviation. We use robust (Huber-White) standard errors.

Table 6 - Relation between Future Return Comovement and Disclosure Similarity around Analyst Brokerage Closures

Panel A - Changes in Disclosure Similarity around Analyst Brokerage Closures

	Disclosure Similarity					
	(1)			(2)		
	Coef.	*t-Stat*		*Coef.*	*t-Stat*	
Treat x Post	-0.076	-3.72	***	-0.058	-2.80	***
Treat	0.242	13.23	***	0.171	9.61	***
Post	0.004	0.96		0.004	0.94	
Same Investor				0.182	22.18	***
Same Industry				0.030	11.26	***
Same Auditor				0.014	6.99	**
Similar MTB				0.017	4.58	***
Similar SIZE				0.021	10.29	***
Similar ROA				0.044	11.43	***
Similar MOM				0.022	5.73	***
Firm *i* x Yearqtr FE	Yes			Yes		
Firm *j* x Yearqtr FE	Yes			Yes		
R-squared	0.69			0.69		
Obs	5,596,033			5,596,033		

Table 6 – Continued

Panel B - Changes in Future Return Comovement and Disclosure Similarity around Analyst Brokerage Closures

	Δ Future Return Comovement								
	Brokerage Closures			*Brokerage Closures*			*Main Sample*		
	(1)			(2)			(3)		
	Coef.	*t-Stat*		*Coef.*	*t-Stat*		*Coef.*	*t-Stat*	
Δ Disclosure Similarity	0.010	4.34	***	0.009	4.05	***	0.001	2.26	**
Δ Earnings Comovement	-0.002	-0.46		-0.004	-0.77		-0.003	-3.86	***
Δ Same Industry	0.043	0.71		0.044	0.72		0.018	1.24	
Δ Same Investor	0.029	4.87	***	0.029	4.84	***	0.103	62.02	***
Δ Same Analyst				-0.030	-0.84		0.006	0.99	
Δ Same Auditor	0.022	1.18		0.022	1.18		0.016	4.07	***
Δ Similar MTB	0.019	3.06	***	0.019	3.05	***	0.001	0.70	
Δ Similar SIZE	-0.005	-0.63		-0.005	-0.63		0.004	3.33	***
Δ Similar ROA	-0.006	-1.05		-0.006	-1.02		0.001	1.12	
Δ Similar MOM	-0.008	-1.54		-0.008	-1.54		-0.005	-5.49	***
R-Squared	0.0003			0.0003			0.0005		
Observations	218,742			218,742			8,296,799		

Notes: Panel A evaluates how disclosure similarity changes when two firms (firms A and B) lose their common analyst as a result of analyst brokerage firm closures. The firm pair that lost common coverage are benchmarked against a randomly selected peer group. For example, firm A is paired with randomly selected peer firms that are not firm B, and vice versa. Each observation in the data represents a pairing of firms (firms *i* and *j*) during a calendar quarter (time *t*). *Treat* is an indicator variable taking the value of one if firm *i* and firm *j* lost a common analyst as a result of analyst brokerage firm closures, and zero otherwise. *Post* is an indicator if the analyst brokerage firm has closed. All other variables are as previously defined. All continuous variables are standardized by subtracting off the mean and dividing by the standard deviation. Standard errors are clustered by firm *i* and year-quarter. Panel B evaluates the relation between future return comovement and disclosure similarity using a changes model. Columns (1) and (2) evaluate the relationship between future return comovement and disclosure similarity in the two quarters after losing a common analyst, as a result of brokerage closures. Column (3) presents benchmark results using the main sample. All variables are as previously defined. All continuous variables are standardized by subtracting off the mean and dividing by the standard deviation. We use robust (Huber-White) standard errors.

Table 7 - Seemingly Unrelated Regression around S&P 500 Inclusion/Exclusion and Analyst Brokerage Closures

Panel A. S&P 500 Inclusion/Exclusion

	Future Return Comovement						*Difference compared using Seemingly Unrelated Regression*			
	(1)			(2)						
	Coef.	*t-Stat*		*Coef.*	*t-Stat*					
Treat x Inclusion	0.107	9.66	***	0.102	9.23	***	0.005	***	4.87%	of Main Effect
Disclosure Similarity				0.102	18.40	***				
Main Effects and Controls	Yes			Yes						
R-Squared	0.08			0.09						
Observations	6,506,065			6,506,065						

Panel B. Analyst Brokerage Firm Closures

	Future Return Comovement						*Difference compared using Seemingly Unrelated Regression*			
	(1)			(2)						
	Coef.	*t-Stat*		*Coef.*	*t-Stat*					
Treat x Post	-0.080	-3.26	***	-0.078	-3.12	***	-0.003	***	3.66%	of Main Effect
Disclosure Similarity				0.055	28.42	***				
Main Effects and Controls	Yes			Yes						
R-Squared	0.11			0.11						
Observations	5,605,174			5,605,174						

Notes: This table evaluates the magnitude change in treatment effects for future return comovement when controlling for disclosure similarity. Panels A and B evaluate the change in treatment effect for S&P 500 inclusion/exclusion, and brokerage closure samples, respectively. *Future Return Comovement* is measured as the Pearson correlation coefficient between a firm's stock return (firm *i*) and a peer firm's stock return (firm *j*) over the 90 days following the later of firm *i* and firm *j*'s disclosure. *Disclosure Similarity* is measured as the cosine-similarity of a firm's 10-K/Q report with that of a peer firm's 10-K/Q report, where peer reports come from the same calendar quarter. For Panel A, *Treat* is an indicator variable taking the value of one if firm *i* is paired with an S&P 500 firm *j* and zero otherwise. *Inclusion* in Panel A is an indicator if firm *i* is in the S&P 500 in period *t*. For Panel B, *Treat* is an indicator variable taking the value of one if firm *i* and firm *j* lost a common analyst as a result of analyst brokerage firm closures, and zero otherwise. *Post* in Panel B is an indicator if the analyst brokerage firm has closed. All other variables are as previously defined. Standard errors are clustered by firm *i* for Seemingly Unrelated Regression.

Table 8 - Relation between Future Return Comovement and Disclosure Similarity around S&P 500 Inclusion

Panel A. S&P 500 Inclusion/Exclusion

	Disclosure Similarity			*Future Return Comovement*		
	(1)			(2)		
	Coef.	*t-Stat*		*Coef.*	*t-Stat*	
Treat x Inclusion	0.019	20.81	***	0.101	105.18	***
Firm *i* x Yearqtr FE		Yes			Yes	
Firm *j* x Yearqtr FE		Yes			Yes	
R-Squared		0.70			0.68	
Observations		6,503,038			6,503,038	

Panel B. Intensity of Treatment Model

	TxI_Future Return Comovement					
	(1)			(2)		
	Coef.	*t-Stat*		*Coef.*	*t-Stat*	
TxI_Disclosure Similarity	0.557	5.56	***	0.324	3.96	***
TxI_Earnings Comovement				0.083	1.73	*
TxI_Same Quasi-Indexer Holdings				0.104	2.80	***
TxI_Same Dedicated Investor Holdings				0.020	1.00	
TxI_Same Transient Investor Holdings				-0.142	-3.50	***
TxI_Same Industry				0.014	0.10	
TxI_Same Analyst				0.565	2.90	***
TxI_Same Auditor				0.086	0.88	
TxI_Similar MTB				0.191	1.13	
TxI_Similar SIZE				0.059	1.39	
TxI_Similar ROA				0.423	2.36	**
TxI_Similar MOM				0.879	3.61	***
TxI_Similar Amihud				-0.039	-0.89	
TxI_Similar Turnover				0.502	2.64	***
R-Squared		0.088			0.322	
Observations		484			484	

Notes: Panel A evaluates how disclosure similarity and future return comovement change upon a firm's inclusion into (exclusion from) the S&P 500. For each firm that entered or exited the S&P 500 during our sample period (firm *i*), we compare their disclosure with other S&P 500 firms and randomly selected non S&P 500 firms (firm *j*) during that time period (*Disclosure Similarity*). *Treat* is an indicator variable taking the value of one if firm *i* is paired with an S&P 500 firm *j* and zero otherwise. *Inclusion* is an indicator if firm *i* is in the S&P 500 in period *t*. *Future Return Comovement* is measured as the Pearson correlation coefficient between a firm's stock returns (firm *i*) and a peer firm's stock return (firm *j*) over the 90 days following the later of firm *i* and firm *j*'s disclosure. All continuous variables are standardized by subtracting off the mean and dividing by the standard deviation. Standard errors are clustered by firm *i* and year-quarter. Panel B evaluates the relation between future return comovement and disclosure similarity using a difference-in-differences changes model following Schoenfeld (2017). This analysis is analogous to an intensity of treatment test, where we compare cross-sectional variation in the treatment effect on *Disclosure Similarity* with that of the treatment effect on *Future Return Comovement*. Each observation in this analysis represents one firm who entered/exited the S&P 500 during our sample period. We estimate models from Panel A by firm *i*, and save the estimated treatment effects (*Treat x Inclusion*). These difference-in-differences estimates at the event-level are demarcated by *TxI* followed by the respective fundamental (e.g., *TxP_Future Return Comovement*). All other variables are as previously defined. Standard errors are adjusted using Huber-White.