

A New Measure of Firm-Group Accounting Closeness ^{*}

Jonathan Ross[†] David Ziebart[‡] Anthony Meder[§]

Abstract

In this study we develop a new measure of firm–group accounting similarity which captures co-movement in accounting fundamentals over time. We calculate the mean of (1) the average pair-wise correlation between earnings and cash flows for a group of firms and (2) the average pair-wise R^2 from regressing firm i 's earnings(cash flows) on firm j 's earnings(cash flows) within-group. We analyze how the four most widely used industrial classification schemes perform in relation to our measure. We document that the within-industry information transfer contagion effect is increasing in our measure. We confirm that within-industry similarity for the average industry varies widely and is the reason why announcement contagion effects are not observed in certain industries. Lastly, we show that the number of analysts following an industry increases in our measure of industry closeness. From an investor's standpoint, the new measure enables them to better discern whether the performance of a group of firms is related over time. From a researcher's standpoint, the new measure enables them to more accurately control for industry effects by capturing the fact that within-industry similarity varies across industries.

Keywords: correlation, accounting comparability, accounting closeness, contagion, information transfer

^{*}We thank Charles Lee for his helpful comments and suggestions which have improved the flow and contribution of our paper. We also gratefully acknowledge the helpful comments and suggestions of John Fellingham, Hong Xie and workshop participants at the University of Kentucky, Rutgers University and Binghamton University.

[†]School of Management, Binghamton University, jross@binghamton.edu, 502-321-5609 (Corresponding author)

[‡]Von Allmen School of Accounting, University of Kentucky

[§]School of Management, Binghamton University

1 Introduction

In this study, we examine the concept of firm similarity. We use the term “closeness” to distinguish our definition from what others may think of when they refer to “comparability”. We define closeness as the extent to which two firm’s accounting fundamentals (e.g. earnings and cash flows) move together over time. We view the financial statements as a function of the underlying economic transactions; the function being the accounting system (e.g. firm-specific application of GAAP) where the accounting function records transactions, aggregates, allocates, and produces financial statements. We measure closeness using cash flows and net income (or earnings). This is the first measure of closeness that is purely accounting-based and it is not difficult to calculate.

Current industrial classification schemes¹ put firms into groups primarily based on similarity in primary revenue generating activities. Whether these classifications lead to groups with reasonable co-movement in accounting measures such as (earnings and cash flows) across time is an empirical question that we answer. Recent research has examined additional means of classification; Hoberg and Phillips (2016) group based upon product descriptions provided in 10-K filings. Lee et al. (2015) uses investor sentiment (proxied by SEC Edgar search timing characteristics) to group firms. We have a different motivation; we develop a measure of within-group accounting similarity for the industrial classification groups. That is, we do not develop a new way to classify firms into groups.

Our measure consists of two parts: C_1 and C_2 . C_1 is the average of all pair-wise Pearson correlations in earnings and cash flows for a group of firms over a given period of time. C_2 is the average of all pair-wise earnings and cash flow correlation matrix determinants² for a group of firms over a given time period (i.e. the mean $1 - R^2$ of the regression of firm i ’s earnings (cash flows) on firm j ’s for all $i \neq j$). Our measure, \mathcal{C} , is the average of C_1 and

¹e.g. Standard Industrial Classification (SIC), Global Industrial Classification (GIC), North American Industrial Classification (NAICS) and Fama-French (FF).

²The correlation matrix determinant captures co-movement in a different way than individual off-diagonal correlations.

$1 - C_2$. We measure closeness over the 25, three-year rolling windows during the 1990-2014 period; with C having a range of -1 to 1. The four main industry classifications all have similar mean C values, but the ranges do vary. Also, for our measure of closeness each scheme produces groups whose mean industry closeness is significantly greater than randomly chosen groups from the population.³

We are primarily motivated from a methodological perspective. Current research uses industry dummy variables to control for industry effects. This inherently assumes that within-industry similarity is constant across industries; the firms in the real estate development industry are just as similar as the firms in the gas utility industry. With our new measure, we find that this assumption is not supported. We find that within-industry similarity varies substantially across industries and time.⁴ Controlling for industry effects with dummy variables is noisy; industry-specific accounting, business conditions for the industry, market sentiment for the industry, and regulatory conditions for the industry are just some of the effects conflated in the dummy variable. If a researcher is focusing on accounting closeness and its effects (such as studies regarding accounting contagion) then our closeness measure will provide a more precise proxy for the accounting effect of each industry. Our measure also provides a way to compare the closeness across industries that a simple dummy variable cannot provide.

Propensity score matching is another common method used in prior research to control for industry effects. Propensity scoring requires matching pairs of firms based on numerous underlying characteristics of interest and this requirement often greatly reduces sample size. Our closeness measure should not limit sample size as propensity scoring often does; it is based on commonly reported accounting information. Although we present our measure as a function of earnings and cash flows, a researcher can easily modify it to include any number of other firm-level characteristics deemed important.

Understanding that within-industry similarity is not constant across industries can lead

³See Table 2.

⁴See Figure 1 & 2.

to a more refined interpretation of the results of prior research studies with industry effects. For example, we replicate the first part of Gleason, Jenkins, and Johnson (2008) to show what effect the violation of this assumption can have on the interpretation of their results. We replicate their finding that a 0.5% statistically significant contagion effect exists around the announcement date of an accounting restatement for the average industry. We show that this effect is significantly larger for closer industries. Specifically, for 8-digit GIC industries with a moderately high value on our measure, we find that a one-percent decrease in the abnormal return for a firm announcing an accounting restatement spills over into a statistically negative cumulative abnormal return for the other firms in that industry over a one-day window. Similar results are also found for levels of our measure when the return window is two or three days. More importantly, perhaps, is that when \mathcal{C} is low no contagion effect is observed. Our measure provides more granularity for closeness and contagion studies; rather than taking all classifications as equally close, our measure can provide ‘closeness thresholds’ above which contagion is expected.

We validate our measure in two contagion effect contexts: earnings announcement contagion and the aforementioned accounting restatement contagion. In both contexts we find that the contagion effect of the announcement on those industry “peer” firms is increasing with our measure of industry closeness. Industries with higher levels of closeness have stronger contagion effects *ceteris paribus*. For example, using a large cross sectional sample of firms over the sample period 1990–2014 we show that, when industry closeness meets a threshold ($\mathcal{C} \geq 0.2217$),⁵ a simple trading strategy that goes long(short) in the peer firms of an industry leader that announces good(bad) news and unwinding the position after 1-3 days earns abnormal returns in the range of 0.44 to 1.08%. These validation exercises also can speak to future research methodology regarding industry effects. Since within-industry similarity varies across industries, we advocate that researchers control for this variation by using our measure of industry closeness rather than a simple industry dummy variable which

⁵The highest decile of closeness.

fails to capture this variation.

The discussion of our analyses and results is organized as follows. Section 2 examines how we operationalize the closeness construct and develop our closeness measure. Section 3 synthesizes the relevant related research. Section 4 examines how the current industrial classification schemes perform in relation to our measure. Our three validation tests of \mathcal{C} are provided in Section 5. We provide a summary of our conclusions in Section 6.

2 The Measure

We view a given firm's accounting system as mapping economic events to financial statements:

$$y_i = f_i(x_i) \tag{1}$$

where x_i are the economic transactions that firm i engaged in during the period, f_i is their accounting function (i.e. essentially their specific application of GAAP, or set of journal entries if you will) and y_i is their set of financial statements.

From equation (1) we see there are two possibilities for measuring the closeness between firms. One approach is to measure the similarity in functions as done in De Franco, Kothari, and Verdi (2011).⁶ The other approach is to measure the similarity of outputs (i.e. similarity in financial statements). Investors cannot observe the accounting functions but they can observe the outputs. Measuring closeness as the degree to which y_i and y_j are similar seems more appropriate from an investor's perspective. If firms release their financial statements at different times, the investor can use the information in the first announcer's financial statements to predict the information that will be later announced by the other firm based on their knowledge of past financial statement similarity, regardless of whether the unobservable f_i and f_j are similar or not.

⁶De Franco, Kothari, and Verdi (2011) provide evidence that their measure is positively related to analyst forecast accuracy and negatively related to analyst forecast dispersion.

The financial statements provide an overall picture of the financial health of a company and provide key indicators which help market participants judge the future performance of the company. Since we consider closeness from an investors perspective, we feel the similarity in outputs takes precedence as investors are less interested in the specifics of the accounting and more interested in the financial and economic outcomes of the firm.

Based on the preceding discussion we measure “closeness” as **the degree of similarity in the outputs of the accounting system** rather than the degree of similarity in what we loosely refer to as the “structure” of the accounting system (i.e. the degree of similarity between f_i and f_j). The outputs of the accounting system should reflect the underlying transactions from which they are derived and thus if two firms have fundamental outputs which move predictably (i.e. exhibit reasonable co-movement), then we view those firms as “close” or similar whether or not they process their respective transactions similarly.

We assume that the summary performance measures (e.g. earnings and cash flows) of a firm capture the performance information available in the financial statements and thus all available information in the underlying transactions. Observing the correlation between these measures across time and between firms will capture closeness as we define it. More specifically, we know that both earnings and cash flows capture changes in assets, liabilities and equity across periods. In this sense, the assumption that earnings and cash flows capture all available performance information in the financial statements is warranted.

Our closeness measure consists of two parts, C_1 and C_2 . Both parts can be functions of whatever fundamental outputs of the accounting system one wishes to use (e.g. earnings, cash flows, total assets, R&D etc.) and any linear combination of those outputs. A benefit of this flexibility is that we can measure closeness of any accounting output of interest. Given the discussion regarding the precedence of earnings and cash flows, we introduce the measure as a function of these two fundamental outputs. C_1 is the mean of the pair-wise Pearson correlations in earnings and cash flows for a given group of firms over a given time period. For example, given an earnings and cash flow time series over a five-year period for each

of four firms, C_1 is the mean of the six pair-wise earnings correlation coefficients and six pair-wise cash flow correlation coefficients. In a given year, earnings and cash flows for a firm diverge due to the accrual process and both contain valuable information independent of the other. The earnings correlation coefficient between a pair of firms may not capture the whole story in terms of fundamental similarity. For example, the firms may exercise different timing in their accruals and this will downwardly bias the earnings correlation coefficient. On the other hand the cash flow correlation coefficient will capture the degree of co-movement in underlying cash flows over time for this pair of firms. Averaging the pair-wise earnings and cash flow correlation coefficients provides a truer measure of the underlying fundamental co-movement than considering each of these separately since it captures both dimensions of accrual earnings and cash flow.⁷

C_2 is the mean of the pair-wise earnings and cash flow 2×2 Pearson correlation matrix determinants for a given group of firms over a given time period. Using the same example as earlier, C_2 is the mean of the six pair-wise earnings correlation matrix determinants and the six pair-wise cash flow correlation matrix determinants. The determinant of a pair-wise earnings 2×2 correlation matrix is a measure of the similarity in earnings between two firms and captures something different than the correlation itself. Correlation captures the extent to which firm j 's earnings change relative to j 's mean when firms i 's earnings change relative to i 's mean. In contrast, the determinant captures that portion of variation in firm i 's earnings that is not captured by variation in firm j 's earnings and is equivalent to $1 - R^2$ where R^2 is the percentage of variation in firm i 's earnings explained by variation in firm j 's earnings in an OLS regression of firm i 's earnings on j 's. Notice that $0 \leq C_2 \leq 1$. The closer C_2 is to zero the closer R^2 is to 1 implying that variation in firm j 's earnings can explain more variation in firm i 's earnings.

In summary, C_1 and C_2 capture co-movement in the underlying accounting fundamentals but in different ways. C_1 captures correlation or the extent to which firm i 's earnings and

⁷Our results are robust to using only earnings or cash flows in the calculation of C_1 .

cash flows change relative to i 's respective means when firm j 's earnings and cash flows change relative to j 's respective means. $1 - C_2$ captures that portion of variation in firm i 's earnings(cash flows) that is explained by variation in firm j 's earnings(cash flows). We do not have a reason to believe why one would care more about one dimension of co-movement versus the other; therefore we simply take the average of C_1 and $1 - C_2$ to form \mathcal{C} .

Notice that C_1 is between negative and positive one while C_2 is between zero and one.⁸ Thus C_1 can tell us something about negative co-movement while C_2 cannot. Finally, as discussed earlier, values of C_2 closer to zero indicate closer groups while values of C_1 closer to one and negative one indicate closer groups. Because of the potential confusion we take $1 - C_2$ when forming the closeness measure in equation (2).

$$\mathcal{C} = \begin{cases} \frac{1}{2} [C_1 + (1 - C_2)] & \text{if } C_1 \geq 0 \\ \frac{1}{2} [C_1 - (1 - C_2)] & \text{if } C_1 < 0 \end{cases} \quad (2)$$

The measure \mathcal{C} is simply the average of C_1 and $1 - C_2$. Note that \mathcal{C} is bounded below by -1 and above by 1. Values of \mathcal{C} close to positive or negative one indicate groups whose earnings and cash flows exhibit high co-movement either in the same direction or opposite directions.⁹

Although we don't report it, we find via simulation analysis that C_1 and C_2 both capture similarity in the movement of accounting fundamentals over time but in different ways. Specifically, we find that C_2 is more sensitive than C_1 to changes in the variance of shocks (randomness) to earnings and cash flows for a given group of firms. By averaging C_1 and $1 - C_2$, our measure captures this differential sensitivity.

The main limitation of our measure is that \mathcal{C} measures the closeness in ex post financial statement information; \mathcal{C} does not control for differences in firm-specific exposure to economic

⁸Based on the properties of the Pearson Correlation and determinants of two-by-two correlation matrices.

⁹Both types of co-movement indicate close groups since either type aids in predicting the fundamentals of the whole group by observing the fundamentals of one or more firms in that group.

transactions.¹⁰ Therefore, one must attempt to control for these differences by first putting firms into groups which have exposure to similar economic transactions before calculating \mathcal{C} . Thus \mathcal{C} is not designed to take a population of firms and optimally delineate them into respectively “close” groups. In contrast, \mathcal{C} is designed to, once given a group of firms (e.g. an SIC 2-digit industry), determine the closeness of those firms based on accounting financial statement co-movement.

3 Related Research

Our measure contributes primarily to two streams of literature. First, prior literature has examined the concept of firm–group closeness/comparability and the extent to which current industrial classification schemes group similar firms. De Franco, Kothari, and Verdi (2011) address the concept of “similarity” in an accounting information sense. Their study provides a concise way to analyze a fundamental concept of accounting information. They define the accounting system of a firm as in equation (1) from the previous section and use regression analysis to estimate the accounting functions of firms and then the comparability of their accounting functions.

Prior to De Franco, Kothari, and Verdi (2011), a few studies have offered measures of similarity (e.g. Bhojraj and Lee (2002), Bhojraj, Lee, and Oler (2003) and Chan, Lakonishok, and Swaminathan (2007)) and tested these measures for firms grouped according to the major classification schemes (SIC, GICS, NAICS and FF). These classification schemes usually group firms based on primary revenue generating business activity with diversified firms grouped by industry. These studies have examined how each of these systems capture “economic relatedness” among firms (i.e. co-movement of returns). A few of these studies further examine the degree of homogeneity among firms in regards to accounting variables

¹⁰For example, certain accounting regulations may affect certain firms differently in how they account for specific economic transactions. This difference in accounting treatment would effect the earnings portion of our ex post closeness measure (not the cash flow portion). Our measure cannot delineate between the effect on ex post co-movement of the difference in accounting for given economic events and the ex ante exposure to different economic events.

(sales, earnings etc.) as a result of being grouped according to the classification systems. Our study also examines the extent to which these schemes group firms whose accounting fundamentals co-move over time. Our two-part measure, \mathcal{C} , is different than the measures in these studies due to the addition of C_2 and the fact that \mathcal{C} is a joint function of earnings and cash flows. Also, we provide analysis regarding which industries rank highest on our closeness measure over the last twenty-five years, how industry closeness changes over time for the four most widely-used industrial classification schemes and how industry closeness varies within-industry over time.

A recent study by Hrazdil and Scott (2013) shows that estimates of discretionary accruals derived using the GICS significantly outperform estimates derived using each of the three alternative industrial classification schemes (SIC, NAICS and Fama-French). They call for future research to consider using the GICS. Our study builds on this study and shows that even within the GICS there is relatively substantial variation in closeness: the firms in some industries are closer to each other (in terms of historical earnings and cash flow co-movement) than the firms in other industries are with each other.

Also, recently several papers have examined new ways to group firms into industries. Most notably, Hoberg and Phillips (2016) use text-based analysis of firm 10Ks to group firms based on their product descriptions. Additionally, Lee et al. (2015) apply a “co-search” algorithm to internet traffic at the SEC’s EDGAR website and develop a method for identifying peer firms which captures the collective sentiment of investors. Where these studies use subjective marketing language in the form of product description or aggregate investor beliefs, we simply develop a measure based on accounting fundamentals by which we can assess a given grouping scheme. One could even use our measure to assess the accounting closeness of firms grouped under these new schemes. This is no different than applying our measure to the SIC or GIC scheme groupings as we do later in Section 4.

Our study also contributes to the information transfer literature which examines the industry contagion share price effects of information releases by one of the firms in that

industry on the other firms in that industry. The literature has studied numerous types of information releases. For example, earnings announcements (Foster (1981), Clinch and Sinclair (1987), Freeman and Tse (1992) and Ramnath (2002)), earnings forecasts (Pownall and Waymire (1989)), unexpected earnings (Han and Wild (1990)), and earnings restatements (Gleason, Jenkins, and Johnson (2008)) have been examined previously. Information transfer effects surrounding retailers monthly sales reports, bank loan-loss reserves, bank failures, bankruptcy filings, dividend initiations, internet hacker attacks, nuclear accidents, negative stock price surprises and share repurchase deregulations have also been documented.¹¹

Two of the validation exercises for our measure involve an accounting news announcement and the resulting information transfer. Our first validation exercise examines the intra-industry contagion effects of an earnings announcement by the “leader” of an industry.

Baginski (1987) examines the information transfer effects of management earnings forecasts. These forecasts change the market’s expectation of earnings for the forecasting firm as well as the peer firms, causing the share prices to move for both. He hypothesizes that this contagion effect should only take place if the forecasting firm is “sufficiently” similar to each of the peer firms. He restricted his analysis to a group of ‘similar’ firms and was able to isolate the contagion effect. However, he did not examine how the contagion effect interacted with the similarity of the group. Two questions come to mind upon reading Baginski (1987). First, how does the contagion effect of a management earnings forecast vary with the similarity between the forecasting firm and the peer firm? Second, how ‘similar’ does the forecasting firm need to be to the peer firm before a contagion effect will exist? The results of our study offer insight into these two questions.

Pyo and Lustgarten (1990) also examine the information transfer effects on peer firms from a management earnings forecast announced by firm i . They document that peer firms experience a statistically significant positive(negative) abnormal return when firm i an-

¹¹See, Docking, Hirschey, and Jones (1997), Aharony and Swary (1983), Bowen, Castanias, and Daley (1983), Olsen and Dietrich (1985), Lang and Stulz (1992), Firth (1996), Ferris, Jayaraman, and Makhija (1997), Ettredge and Richardson (2003), Akhigbe, Madura, and Martin (2015) and Chang, Lai, and Yu (2005).

nounces good(bad) news (proxied for by the abnormal return of firm i on their announcement date) and this contagion effect is increasing in the pair-wise historical earnings correlation between firm i and its peers.

Our paper contributes to Pyo and Lustgarten (1990) in several ways. First, we examine the information transfer effects on firm j of an *earnings* announcement by firm i rather than a manager's forecast of earnings. Earnings announcements are mandatory and highly anticipated whereas management forecasts of earnings are voluntary and based on expectations which may or may not be realized. Thus earnings announcements provide a stronger test of an information transfer effect than management forecasts of earnings. Second, our measure of closeness considers not only historical correlation (C_1) but also percentage of variation explained (C_2). Furthermore, the inputs in our measure are not only earnings but also cash flows; capturing both accruals and cash flows. Finally, our sample consists of 1,421 firm earnings announcements paired with 35,324 firm-year observations and our time period covers 25 years (1990–2014) and thus provides a more extensive test of the relationship between contagion effects and firm closeness.

Freeman and Tse (1992) examine the information transfer effects on firm j (late announcer) of an earnings announcement by firm i (early announcer). If firm i is the first announcer then firm j is the 2nd–5th announcer of earnings in a particular industry for a particular quarter and so forth. They document statistically positive contagion effects on firm j if firm i is the first announcer than if firm i is the 2nd announcer and so forth. The contagion effect of a 2nd–4th announcer on those later announcers (3rd–5th) share prices is basically zero. Furthermore, these contagion effects are stronger the greater the pair-wise historical earnings (and sales) correlation between the early announcers and late/non-announcers. Overall they conclude that “earnings-based information transfers of general industry-wide trends are typically small (relative to firm-specific information) and may be limited to industries with strong earnings co-movement.”

Our paper contributes to Freeman and Tse (1992) in the same ways we contribute to

Pyo and Lustgarten (1990). Additionally, we also provide insight into the level of closeness a GIC 6-digit industry¹² must attain for a 1-day, 2-day and 3-day contagion effect to be present (0.06, 0.09, 0.11 respectively). Since we document the average GIC 6-digit industry closeness level is $\mathcal{C} \approx 0.128$, we would conclude differently from Freeman and Tse (1992) and suggest that industry closeness need not be very high, in fact less than average, for a statistically significant contagion effect to exist.

A further contribution our paper makes to Pyo and Lustgarten (1990), Freeman and Tse (1992) and the rest of the contagion literature is we provide detailed information on the magnitude of the contagion effect for those earnings announcements in industries ranking in the highest decile of closeness ($\mathcal{C} > 0.2217$, e.g. Gas Utilities, Water Utilities, Beverages). A trading strategy which goes short(long) in those industries in the highest decile of closeness after bad(good) earnings news announced by the leader earns 0.46-1.08% abnormal returns over 1-3 day windows. To get the 1.08%, for example, one should sell short those peer firms in the highest decile of closeness when leaders announce bad news and buy back after 1 trading day.¹³

We also contribute to Gleason, Jenkins, and Johnson (2008) by showing that the contagion effect of an earnings restatement is also increasing in the closeness of the industry in which the announcing firm resides.

Finally, we contribute to the analyst coverage (Bhushan (1989), Barth, Kasznik, and McNichols (2001)) literature by documenting that the number of analysts which issue an earnings per share forecast for firm i is increasing in the closeness between firm i and the other firms in firm i 's industry.

¹²GIC 6-digit industries are at the same level of granularity as SIC 2-digit industries.

¹³See Section 5 for the criteria we used in determining the "leader" of a given industry in a given year and a more detailed discussion on the trading strategy results.

4 Industrial Classification Schemes and Industry Closeness

We evaluate the ability of the Standard Industrial Classification two-digit (SIC2), four-digit (SIC4), North American Industrial Classification System three-digit (NAICS3), Fama-French (FF) and the Global Industrial Classification six-digit (GIC6) to capture the similarity of accounting fundamentals within their grouping using our measure, \mathcal{C} . Our sample period is 1990-2014 and we measure \mathcal{C} as of the end of a particular industry-year using each firm's 12 previous¹⁴ quarterly income before extraordinary items and operating cash flows for all 12/31 fiscal year-end firms.¹⁵ Industry-years are retained if they have at least 5 firms with the required historical data.

Table 1 reports the pooled mean of \mathcal{C} across the industry years as well as the pooled mean number of firms per industry-year for each of the schemes over our sample time period. Table 1 also reports the number of industries represented as a fraction of the total number of industries per scheme. The schemes are ranked in terms of \mathcal{C} and t-stats from one-tailed hypothesis tests regarding the difference between the GIC6 average \mathcal{C} and the other schemes are reported below each scheme's mean \mathcal{C} . The schemes have insignificantly different average industry accounting fundamental closeness. This is interesting in light of Bhojraj, Lee, and Oler (2003) who find that the GIC scheme groups firms into industries with higher average **stock return co-movement** than the other schemes.

Furthermore, industry closeness displays considerable variation across time within-industry and average industry closeness varies across time as well. This variation is similar for each scheme.¹⁶ To illustrate variation across time within-industry, Figure 1 plots the closeness

¹⁴Results are robust to using 8 or 16 previous quarterly numbers as well as using 5 previous annual numbers. Also, although the GIC, NAICS and FF schemes came along before the start of our sample time period, COMPUSTAT back fills the firm industry codes for these schemes.

¹⁵Co-movement in earnings and cash flows needs to be measured over the same calendar time period. Otherwise there will be some macro-level economic events which some firms experience that other firms don't experience. This will add noise to the closeness measures. 12/31 fiscal year-end firms comprise 67% of the firm-years in COMPUSTAT.

¹⁶We leave out the tables for each scheme displaying industry-year closeness for our entire sample time period

for five of the GIC6 industries over our sample time period. We ranked the industries on average \mathcal{C} over our sample time period and randomly chose one GIC6 industry from each of the 5 resulting quintiles. Notice the variation over time for these five GIC6 industries on \mathcal{C} . To illustrate variation in average industry closeness over time Figure 2 plots the mean closeness across all GIC6 industries for each year. Notice how mean industry closeness varies over time, peaking in 2009 when the financial crises affected all industries leading to higher co-movement in fundamentals.

Figure 3 plots the mean closeness¹⁷ of each GIC6 industry over our sample time period. Notice those industries that tend to have higher closeness. The average closeness of the gas and water utility industries (GIC 551020 and GIC 551040) is 0.475 and 0.284 respectively. The utilities industries are highly regulated. For a given economic event, there is less firm-specific discretion in how utility firms should account for that event. Thus, a given economic event affects the utility firms more similarly than firms in a different industry. The higher levels illustrated in Figure 3 support the assertion that \mathcal{C} captures accounting fundamental co-movement and these utility industries are closer than other industries.

Next we provide another validation¹⁸ that our measure captures the similarity construct. Since the industrial classification schemes group firms by similarity in primary revenue generating activity and these firms likely face similar customer demographics, market conditions etc., firms in the same industry should exhibit higher earnings and cash flow co-movement than firms chosen randomly across all the industries. As of the end of each year in our sample time period we randomly select (with replacement) 1000 same-size groups of firms as the average industry sizes of each of the schemes in our sample respectively and calculate the mean \mathcal{C} of the 1000 groups.

Table 2 illustrates the results averaged across the years. Note that the schemes indeed group firms which exhibit statistically significantly higher accounting fundamental co-movement than a random grouping of the firms. In no year was the mean \mathcal{C} of the random

due to space constraints. These tables are available upon request.

¹⁷Mean of the 25 \mathcal{C} values over 1990–2014 for each industry.

groups higher than the corresponding industry classification scheme mean \mathcal{C} . This is consistent with intuition and provides evidence regarding the ability of \mathcal{C} to capture co-movement.

5 Industry Closeness and Intra-industry Information Transfer

5.1 Closeness and Earnings Announcement Information Transfer

Prior information transfer studies¹⁸ have found empirically that good(bad) news announcements by one firm in an industry tend to lead to price increases(decreases) for peer firms. We hypothesize this effect in our study. Additionally, we hypothesize the contagion effect will increase as the accounting co-movement of the firms in the group increases. In industries where earnings and cash flow co-movement is higher we expect a good(bad) news leader earnings announcement will lead to a greater positive(negative) share price contagion effect compared to the same news announcement in an industry with lower earnings and cash flow co-movement. When a leader announces good(bad) news earnings, investors will expect that the peers will also announce good(bad) news if the historical earnings and cash flow co-movement of the group is high (assuming this co-movement persists). If the historical earnings and cash flow co-movement of a group is low then the leader's announcement doesn't provide much information about the peer firm's earnings and thus there should be little movement in peer firm's prices, *ceteris paribus*. Formally, our two hypotheses stated in alternative form, are as follows.

H1: A good(bad) news earnings announcement by the leader of an industry will lead to a statistically positive(negative) share price movement for peer firms in that industry on, and immediately subsequent to, the leader's earnings announcement date, *ceteris paribus*.

¹⁸e.g. See Pyo and Lustgarten (1990), Freeman and Tse (1992) and Gleason, Jenkins, and Johnson (2008)

H2: The closeness of a given industry interacts with the “news” contained in that industry–leader’s earnings announcement to increase the contagion effect in that industry, ceteris paribus.

To test hypotheses one and two we use a sample period from 1990–2014. For each year, we search the CRSP/COMPUSTAT merged file for firms which have earnings and cash flow data for the 12 previous quarters,¹⁹ report date of quarterly earnings for the 4th quarter of that year, fiscal year–end and GIC 6-digit code.²⁰ We then merge this sample with CRSP to obtain the price and number of shares outstanding as of the end of that year. Next, we delete all non 12/31 fiscal year–end firms²¹ and then, for each GIC6 industry, identify a “leader” for that particular year. The leader is the firm with the highest market value as of the end of that year.²² A check of several industries reveals that this method of identifying the “leader” does identify the leader that many would likely consider the most influential firm. For example, identifying the leader of the Oil & Gas industry this way identifies Exxon Mobil as the leader in most years. We find that the firm with the highest market value in an industry tends to be the first announcer of earnings for that quarter.

We delete any “peer” firm in an industry which announces earnings before the leader and within five days after the leader. This allows us to better isolate the portion of share price movement of the peers due solely to the leader’s earnings announcement.

We then calculate the closeness, \mathcal{C}_j , of industry j as of the end of year t using equation (2) with the earnings and cash flow data for each firm in industry j over the 12 previous quarters.

¹⁹Technically our sample period is from 1988–2014 but we only look at contagion effects resulting from earnings announcements between 1990–2014. We need the extra years because the closeness is calculated using earnings and cash flow data over the previous 12 quarters for each year between 1990–2014. Also, quarterly cash flow data (OANCFY) is not reported in the merged database until 1987.

²⁰We use earnings before extraordinary items (IBQ) and net cash flows from operating activities (OANCFY) to measure earnings and cash flows respectively. To calculate operating cash flows in a quarter we take the quarterly change in OANCFY.

²¹Co-movement in earnings and cash flows needs to be measured over the same calendar time period. Otherwise there will be some macro–level economic events which some firms experience that other firms don’t experience. This will add noise to the closeness measures.

²²Our results are robust although not as significant when we use “largest assets” to delineate the leader.

Our final sample consists of 1,421 industry leader annual earnings announcements paired with 35,324 peer firm-year observations representing 6,049 unique firms and all 67 GIC 6-digit industries. The mean number of GIC 6-digit industries represented per year is 59.²³ The mean(median) number of peer firms per industry leader earnings announcement is 24.86(15).

Our sample is comprised of leaders and peer firms from all 67 GIC industries and thus is a fair representation of the population of firms over the same time period. Table 3 reports descriptive statistics for our sample versus the COMPUSTAT population (in parentheses) of firms with annual earnings and asset data over the same time period. Sixty-seven percent of the firms in this population are 12/31 fiscal year-end firms. Although we only use a subsample of these 12/31 firms due to data requirements for the variables, the pooled descriptive statistics of our sample and the population are relatively similar.

We estimate each firm's beta using the standard market model²⁴ Firm daily returns for estimating the market model are obtained from CRSP. Daily market return and risk-free rate are obtained from Ken French's website.²⁵ We use the abnormal return of the leader on the announcement date to proxy for the "news" of their announcement. Positive abnormal returns imply "good news" and negative abnormal returns imply "bad news".²⁶

We estimate the cumulative abnormal returns of the peer firms over windows $w = [0, 1], [0, 2]$ and $[0, 3]$ respectively. Zero represents the leader's earnings announcement date.

Table 4 displays the mean and median contagion effect (peer cumulative abnormal returns) over each of the three windows for the sample in aggregate. The sample is broken down by "good" versus "bad" news. Of the 1,421 earnings announcements, 731(51.4%) were good news while 690(48.6%) were bad news. The average news across all announcements was an insignificant 0.15%. Since slightly more announcements were good than bad, the good news peer firm sample was larger (18,259 versus 17,065) as seen in Table 4. Notice that

²³There are 67, GIC 6-digit industries. See <https://www.msci.com/gics>.

²⁴Using the Fama and French (1993) 3-factor model produces similar results.

²⁵<http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/>

²⁶We did not use analysts forecasts to evaluate the "news" due to limited data availability over our sample period. I/B/E/S has much more limited data coverage than CRSP (especially over the early years in our sample period) and we wanted to maintain a representative sample.

the contagion effect is always consistent in direction to the announcement. That is, when the news is good(bad) the mean peer firm return is positive(negative). Also, on average, the good news is more positive than the bad news is negative. Therefore, the contagion effect of good news announcements is greater than the contagion effect of bad news announcements across the three windows. Thus Table 4 provides evidence in favor of H1.²⁷ The evidence in Table 4 is also consistent with prior research regarding the contagion effects of earnings announcements (e.g. Freeman and Tse (1992)). The larger sample size and longer time period of our study should enhance the generalizability of our results.

Our conclusions are valid to the extent that our results are due to the closeness and contagion effects resulting from the accounting measures and not other economic events specific to the peer firms. We mitigate this risk with sample selection criteria where we eliminate firms which announce earnings within a short time-period of the leader's announcement. We are unaware of any other systematic peer firm-specific events which would lead to the peer returns moving exactly as we hypothesize.

Table 5 reports the mean contagion effect both for good news and bad news earnings announcements for the 1st and 10th decile of industry closeness.²⁸ For example, a decile 1 peer is a firm in an industry whose closeness (measured by \mathcal{C}) is less than or equal to the 10th percentile of closeness across all industries. This level is $\mathcal{C} = 0.0699$ from Table 5. Notice the difference in contagion effects on those peers located in the extreme groups of industry closeness is statistically and economically different than zero over each short window. Also note that abnormal returns as high as 1.08% could be earned (before transaction costs) by selling short those peer firms in the highest decile of closeness when leaders announce bad news and buying back after 1 trading day. Finally, from Table 5, abnormal returns from a buy strategy could be as high as 0.75% if one bought those peer firms in the highest decile of closeness when the leaders announce good news and sold after 1 trading day. Note that

²⁷A simple regression of the peer cumulative abnormal returns on the news of the leaders pooled over all three windows produces a coefficient of 0.0873(p-value = 0.0000 and $R^2 = 0.28\%$.)

²⁸The results are similar if comparing the contagion effects between the 1st and 5th quintiles of industry closeness.

the lowest abnormal return one could make on the highest decile of closeness for a buy–sell or sell short–buy back strategy is 0.46%.

H2 hypothesizes that the closeness of the industry moderates the contagion effect of a leader announcing good or bad news on industry peers. In the simple descriptive analysis in Table 5 we turn two dials and observe an increasing contagion effect as closeness increases. The magnitude of the leader news is not held constant however between the two closeness deciles. It seems reasonable that for a fixed level of industry closeness, the peer firm returns will be increasing(decreasing) in the positiveness(negativeness) of the “good”(“bad”) news announced by the leader. Evidence consistent with this was provided in Table 2 where we document an overall contagion effect across all industries regardless of their closeness. Thus, we cannot safely conclude from Table 5 that the different contagion effects experienced by those peer firms across deciles of closeness is directly a result of the different closeness levels of their industries. To test H2 more directly we run the following pooled regression with our sample of 35,324 firm–years for each of the three windows respectively.

$$CAR_{iw} = \gamma_1 + \gamma_2 \mathcal{C}_j + \gamma_3 News_j + \gamma_4 \mathcal{C}_j * News_j + \epsilon_{iw} \quad (3)$$

CAR_{iw} is the cumulative abnormal return for peer firm i over window w around the leader’s earnings announcement date, \mathcal{C}_j is the closeness of industry j using the measure in equation (2) as of the earnings announcement date where j represents the industry in which firm i resides and $News_j$ is the abnormal return of the leader for industry j on their earnings announcement date. H2 predicts that the coefficient on the interaction between closeness and news, γ_4 , will be statistically greater than zero. Precisely, γ_4 shows how the effect of $News_j$ on R_{jw} (the contagion effect) changes for a one percent increase in \mathcal{C}_j . A statistically positive γ_4 implies that the contagion effect is increasing with the closeness of the industry.

Table 6 Panel A reports the results from a pooled estimate of regression equation (3) with standard errors clustered at the firm and year level²⁹ for each of the three windows.

²⁹Petersen (2009)

The results in Table 6 Panel A are consistent with H2. Specifically, the contagion effect of an industry-leader earnings announcement on those industry peer firms is increasing in the closeness of the industry in which the leader resides. This interaction holds over a one, two and three-day window subsequent to an industry-leader earnings announcement.³⁰

Table 6 Panel A also provides evidence regarding the level that closeness needs to be for a contagion effect to exist. For a one day window, note that the coefficient on *News* is $-0.058 + 0.939\mathcal{C}$. Thus for a one percent increase in the abnormal return of the leader on the date they announce earnings, the abnormal return of peer i increases by $-0.058 + 0.939\mathcal{C}$. Simply set this coefficient equal to zero and solve for \mathcal{C} to determine the level that closeness needs to be for a contagion effect to exist: $-0.058 + 0.939\mathcal{C} = 0 \implies \mathcal{C} \approx 0.062$. Therefore if $\mathcal{C} > 0.062$ a contagion effect will be observed over a one day window subsequent to an industry-leader earnings announcement. Since the average GIC6 industry in our sample has $\mathcal{C} = 0.128$,³¹ a contagion effect will exist for the average industry-leader earnings announcement. Repeat this process to find that \mathcal{C} needs to be at least 0.09(0.11) for a two(three) day contagion effect to exist. The results from Table 6 thus provide insight into how prior researchers could find contagion effects without considering the closeness of the industry. Our findings suggest that the industry dummy method that is dominant in the literature returns significant findings because mean closeness ($\mathcal{C} = 0.128$) is high enough for a contagion effect to exist, on average. Our results extend those findings by showing the degree of closeness matters and that as the length of the windows increases, the closeness threshold for contagion to exist also increases.

Using C_1 or $1 - C_2$ alone in regression equation (3) instead of \mathcal{C} (the mean of these two statistics) produces weaker results. Table 6 Panels B and C report the results from using these two statistics respectively. Notice the results for $1 - C_2$ are stronger than for C_1 . We

³⁰The Table 6 result is robust to the following specifications: (1) Using quarterly data over the previous 16 months (instead of 12 months) to measure \mathcal{C} for each industry-year; (2) Using annual accounting data over the previous 5 years to measure \mathcal{C} instead of quarterly data to measure \mathcal{C} for each industry-year; (3) Using net income and total cash flows instead of earnings before extraordinary items and operating cash flows as inputs in calculating \mathcal{C} ; (4) Measuring industry closeness using one of the other industrial classification schemes such as SIC2, NAICS3 or FF.

³¹Recall Figure 3.

attribute this to the fact that $1 - C_2$ is more sensitive to shocks to earnings and cash flows than C_1 over time.³² This highlights the notion that $1 - C_2$ may be a more precise measure of closeness than C_1 . Collectively, the results from Table 6 imply that averaging C_1 and $1 - C_2$ to form \mathcal{C} leads to larger and more statistically significant interaction coefficients with higher explanatory power. Thus the contagion effect is more sensitive to closeness measured using \mathcal{C} than using either C_1 or $1 - C_2$ alone.³³ This result suggests that both accruals and cash flow provide information that is not redundant and illustrates the usefulness of both.

In summary, the results from our earnings announcement contagion effect test help bolster the construct validity of our measure since we find that the contagion effect is increasing in our measure of closeness; consistent with our hypothesis.

5.2 Closeness and Accounting Restatement Information Transfer

Gleason, Jenkins, and Johnson (2008) hypothesize and find that when one firm in an industry announces an accounting restatement, that firm experiences an average abnormal return of -19.8% over the two day window $[-1,+1]$ surrounding the announcement. Furthermore, those peer firms in the same industry, on average, experience a statistically significant negative return of -0.5% over that same window. We first replicate their study and find qualitatively similar results. Then we extend their study to show that the share price contagion effect of an accounting restatement is increasing with the closeness announcing firm's industry. Formally, our third hypothesis is stated in alternative form below.

H3: The closeness of a given industry interacts with the “news” contained in a firm’s accounting restatement to exacerbate the contagion effect on the peer firms in that industry, ceteris paribus.

As in Gleason, Jenkins, and Johnson (2008) (henceforth GL [2008]), the initial sample of 919 restatement events comes from the GAO (2003) and comprises nearly all of public

³²We determine this in unreported simulation analysis.

³³As an alternative to the equation (??) specification, we replaced cumulative abnormal returns with raw returns and included the cumulative market return as a control variable. The results were very similar to our main specification results from Table 6.

company restatements from January 1, 1997 to June 30, 2002. Contrasting with GL (2008), we did not augment the sample with other restatement events from the Wall Street Journal.³⁴ We followed their sample selection criteria and discarded events if the CRSP announcement period common stock returns were unavailable or the restating firm lacked an industry peer group comprised of at least five firms.

To enhance the ability to detect restatement-induced stock price contagion, we eliminated events where the restatements were initially announced as part of the firm's routine quarterly earnings press releases. We did not want to confound the contagion effects of earnings releases with those of the restatement events themselves. The final sample consisted of 161 restatements representing virtually every economic sector.³⁵

Non-restating peer firms were identified through COMPUSTAT using eight-digit GICS codes.³⁶ We retained those peer firms which satisfied the following criteria: (1) announcement period $[-1,+1]$ stock returns are available on CRSP; (2) the pre-announcement stock price is at least \$5.00; and (3) the peer firm has not announced an accounting restatement within the preceding 24 months. This process produced a peer group sample comprised of 3,251 peer firms or about 20 firms for each restatement event. Our final sample of peer firms is significantly smaller than GL (2008) due to our requirement for measuring closeness. In addition to the GL (2008) criteria, for each restatement event we required at least 12 previous quarters of earnings and cash flow data for each peer firm for calculating \mathcal{C} for the 8-digit GIC industry containing the restating firm.

As in Table 1 Panel A of GL (2008), we calculated the cumulative abnormal return over each of the following announcement windows both for the restating firm and for the peer

³⁴See footnote 10 in GL (2008). They add an additional 185 restatement events identified through their extensive review of approximately 2,200 pertinent WSJ articles. These 185 events make up less than 20% of their initial sample (before applying their sample criteria cuts) and there is no reason to believe these events are qualitatively different enough from the ones contained in the GAO database to justify their inclusion to maintain a "representative" sample. Furthermore, this process requires considerable amount of time and effort since 2,200 WSJ articles must be read carefully. Consequently we do not augment our sample following GL (2008).

³⁵Nearly every eight-digit GICS code was represented.

³⁶We follow GL (2008) in this respect. Our results are entirely robust to using six-digit GIC codes or two-digit SIC codes.

firms; $[-1,+1]$, $[-10,-2]$ and $[+2,+10]$. Table 7 displays the results from our replication along with the results from Table 1 Panel A of GL (2008) in parentheses for comparison. The replication results in Table 7 are qualitatively similar to GL (2008) despite the fact that we use a smaller subset of the GL (2008) sample due to data requirements. Note that the contagion effect on those peers is higher in our replication than in GL (2008). Also, the mean(median) closeness of the GIC eight-digit industries corresponding to the 161 restatement events was 0.140(0.132).

We tested the relationship between the closeness of the industry and the bad news announced by the restating firm. According to H3, we expect the peer firms share prices to fall more when the industry of the announcing firm is closer. We test this hypothesis using regression equation (3). Consistent with GL (2008), we measure the “News” using the three-day CAR of the announcing firm over the window $[-1,+1]$ rather than the one-day CAR over the announcement day. Our results are provided in Table 8 where we perform a pooled estimate of regression equation (3) with standard errors clustered at the firm and year level for each of the three windows.

The evidence in Table 8 is consistent with our hypothesis. The positive coefficients on the interaction between closeness and “News” provide evidence of a contagion effect that is statistically increasing with the closeness of the group in which the restating firm resides. The coefficients on the interaction in Table 8 are much larger in magnitude than the coefficients on the interaction in Table 6 Panel A for the two and three-day windows. The contagion effect from an accounting restatement on those peer firms is increasing more with the closeness of the restating firm’s industry than the contagion effect is for those leaders announcing earnings. Earnings announcements are highly anticipated and forecasted each quarter. Restatements on the other hand are more of a surprise. Evidently, the market feels the restatement announcement tells more about those peer firms as closeness increases than an earnings announcement.

Finally, using the coefficients in Table 8 and equation (3), the coefficient on $News_j$

is $(-0.192 + 1.603\mathcal{C})$ for a one-day window. To find a contagion effect it must be that $-0.192 + 1.603\mathcal{C} > 0$ or $\mathcal{C} > 0.120$. For a two-day(three-day) contagion effect to exist simply follow the same procedure; \mathcal{C} needs to be $> 0.132(0.134)$. Since the mean closeness in our sample of GIC8 industries is $\mathcal{C} = 0.140$ over this time period, a contagion effect will exist in the average industry.

To summarize, Table 7 replicates GL (2008) and provides evidence of a negative contagion effect from a firm announcing an accounting restatement on its industry peers. Table 8 provides evidence that this effect is increasing with the closeness of the industry in question where the closeness is measured using \mathcal{C} . Thus the results in Table 8 help to validate \mathcal{C} as a measure of accounting closeness.

5.3 Industry Closeness and Analyst Forecast Coverage

We provide another validation of our closeness measure in a non-news setting by testing whether the number of analysts issuing an earnings per share forecast is increasing in our measure. Veldkamp (2006) models analysts as information intermediaries in a setting where information is a non-rival good. Under the assumption that information is costly to acquire, she shows investors have higher demand for information that can tell them about many different firm's prospects. The closer the firms in a given industry are, the higher investor demand is for information about any one of those firms as the information can simultaneously tell them about the other firms.

Thus if analysts are suppliers of this information, the more a firm's prior fundamentals comove with their peers the more information analysts will want to provide about firm i . This leads to the prediction that more analysts will cover firms which reside in industries with higher historical earnings and cash flow co-movement. Investors can utilize the forecasts made by the analysts for a firm or firms in the industry to infer earnings performance of the firms for which the investor did not acquire information directly. The implicit assumption is that the historical co-movement continues to persist.

Prior literature (e.g. Bhushan (1989), O'Brien and Bhushan (1990), Brennan and Hughes (1991), Lang and Lundholm (1996), Barth, Kasznik, and McNichols (2001)) has also documented other determinants of analyst coverage. Other variables that have been shown to be statistically related to analyst coverage include size, book-to-market ratio, trading volume, whether or not a firm issues new debt or equity, R&D expense, depreciation expense, the ability of firm i 's past earnings to predict firm i 's future earnings, volatility in earnings and volatility in returns. We control for these other variables in our analysis. Our fourth hypothesis, stated in alternative form, is:

H4: The number of analysts issuing an earnings per share forecast for firm i is increasing in the closeness of the industry in which firm i resides, *ceteris paribus*.³⁷

We begin with the sample of firms from our Section 5.1 earnings announcement contagion test. We test H4 with the following OLS regression over the sample time period 1990–2014.³⁸

$$Coverage_{it} = \beta_1 + \beta_2 \mathcal{C}_{it-1} + \gamma Controls_{it-1} + \epsilon_{it} \quad (4)$$

where $Coverage_{it}$ is the natural logarithm of the number of analysts issuing an earnings per share forecast for firm i for period t using analyst's earnings forecast data provided by I/B/E/S. \mathcal{C}_{it-1} is the closeness of the industry in which firm i resides as of the end of period $t - 1$ calculated using the five previous annual earnings and cash flow realizations for each firm in that industry. $Controls_{it-1}$ is a set of control variables described earlier using period $t - 1$ data.³⁹

Table 9 summarizes the results from regression equation (4) estimated as a panel where the standard errors are clustered at the firm and year level. Table 9 shows that the analyst coverage of a firm is statistically increasing with the closeness of the industry in which that

³⁷Again, we care about the closeness of the industry, not just the pair-wise closeness of firm i with each of its peers for the same reason as discussed earlier after H1 and H2.

³⁸This is the same regression as regression equation (10) in De Franco, Kothari, and Verdi (2011) but we use our measure of closeness instead of their measure of comparability. The reader should compare our Table 9 analyst coverage results to their Table 5 results.

³⁹These control variables are defined in Table 10.

firm resides. Given that the average firm in our sample is covered by 7.89 analysts, a 1% increase in industry closeness for the average firm leads to a 0.801% increase in the natural logarithm of 7.89 analysts. This equates to an increase of 0.13 ($=7.89^{1.00801} - 7.89$) analysts or a 1.66% relative increase in the number of analysts which is a modest increase economically.⁴⁰

In summary, Table 9 provides evidence in favor of H4 and thus we conclude that the number of analysts covering a firm increases with the historical accounting fundamental co-movement (closeness) of the industry in which that firm resides. Thus the results in Table 9 help to validate our measure of accounting closeness.

6 Conclusion

In this study we develop a new measure of accounting closeness, \mathcal{C} , which captures the extent to which the accounting fundamentals (earnings and cash flows) for a group of firms move similarly over time. Current research practice usually controls for industry effects using industry dummy variables. This inherently assumes that within-industry similarity is constant across industries. With our new measure, we find that this assumption is not warranted. We find that within-industry similarity varies substantially across industries and time.

We also validate our measure in several contexts. First, using a 12-quarter rolling window sample time period from 1988–2014 we calculate the mean GIC 6-digit industry closeness using our measure and find that this mean is higher than the mean of our measure calculated for randomly chosen groups of firms (from the same sample) over the same time period. This helps to validate our measure because it is intuitive that firms which have the same primary revenue generating activity (the criteria for industry membership under the GIC for example) likely face similar customer demographics, regulations, costs etc. which should lead to these

⁴⁰We also estimate equation (3) where \mathcal{C} is the average of all pair-wise closeness values between peer i and each of its peers instead of the usual specification which considers all possible pairs of firms in firm i 's industry. The coefficient on \mathcal{C} in this specification was also significantly greater than zero but not as significant ($\beta_2 = 0.332$, t-stat = 4.73).

firm's earnings and cash flows having higher co-movement than firms which have vastly different primary revenue generating activities.

Next, we extend the accounting announcement information transfer literature by showing that the contagion effect of the leader of an industry announcing good or bad earnings news is increasing with the closeness of that industry. Also, we extend Gleason, Jenkins, and Johnson (2008) by showing that the contagion effect of a firm announcing an accounting restatement is increasing with our measure of industry closeness. In industries where closeness is quite low, no contagion effect is observed.

Finally, we extend the analyst coverage literature by hypothesizing and documenting that the number of analysts covering a firm is increasing with the closeness of the industry in which that firm resides.

The results in this study provide researchers with another measure to use in assessing the accounting closeness of a group of firms. Our results also imply that using industry dummy variables to control for differences across industry is limited. When a more contextual measure of industry similarity is desired, we have provided a more informative closeness measure with this study. Accounting fundamental co-movement does vary substantially across industries and seems to be reasonably persistent (doesn't vary extensively over time within-industry). Thus, the closeness of the industry in question should also be considered.

References

- Aharony, J. and I. Swary (1983). Contagion effects of bank failures: Evidence from capital markets. *Journal of Business*, 305–322.
- Akhigbe, A., J. Madura, and A. D. Martin (2015). Intra-industry effects of negative stock price surprises. *Review of Quantitative Finance and Accounting* 45(3), 541–559.
- Baginski, S. P. (1987). Intraindustry information transfers associated with management forecasts of earnings. *Journal of Accounting Research*, 196–216.
- Barth, M. E., R. Kasznik, and M. F. McNichols (2001). Analyst coverage and intangible assets. *Journal of Accounting Research*, 1–34.
- Bhojraj, S. and C. Lee (2002). Who is my peer? a valuation-based approach to the selection of comparable firms. *Journal of Accounting Research* 40(2), 407–439.
- Bhojraj, S., C. Lee, and D. K. Oler (2003). What’s my line? a comparison of industry classification schemes for capital market research. *Journal of Accounting Research* 41(5), 745–774.
- Bhushan, R. (1989). Firm characteristics and analyst following. *Journal of Accounting and Economics* 11(2), 255–274.
- Bowen, R. M., R. P. Castanias, and L. A. Daley (1983). Intra-industry effects of the accident at three mile island. *Journal of Financial and Quantitative Analysis* 18(01), 87–111.
- Brennan, M. J. and P. J. Hughes (1991). Stock prices and the supply of information. *The Journal of Finance* 46(5), 1665–1691.
- Chan, L. K., J. Lakonishok, and B. Swaminathan (2007). Industry classifications and return comovement. *Financial Analysts Journal* 63(6), 56–70.
- Chang, S.-C., J.-H. Lai, and C.-H. Yu (2005). The intra-industry effect of share repurchase deregulation: Evidence from taiwan. *Review of Pacific Basin Financial Markets and Policies* 8(02), 251–277.
- Clinch, G. J. and N. A. Sinclair (1987). Intra-industry information releases: A recursive systems approach. *Journal of Accounting and Economics* 9(1), 89–106.
- De Franco, G., S. Kothari, and R. S. Verdi (2011). The benefits of financial statement comparability. *Journal of Accounting Research* 49(4), 895–931.
- Docking, D. S., M. Hirschey, and E. Jones (1997). Information and contagion effects of bank loan-loss reserve announcements. *Journal of Financial Economics* 43(2), 219–239.
- Ettredge, M. L. and V. J. Richardson (2003). Information transfer among internet firms: the case of hacker attacks. *Journal of Information Systems* 17(2), 71–82.
- Fama, E. F. and K. R. French (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33(1), 3–56.

- Ferris, S. P., N. Jayaraman, and A. K. Makhija (1997). The response of competitors to announcements of bankruptcy: An empirical examination of contagion and competitive effects. *Journal of Corporate Finance* 3(4), 367–395.
- Firth, M. (1996). Dividend changes, abnormal returns, and intra-Industry firm valuations. *Journal of Financial and Quantitative Analysis* 31(02), 189–211.
- Foster, G. (1981). Intra-industry information transfers associated with earnings releases. *Journal of Accounting and Economics* 3(3), 201–232.
- Freeman, R. and S. Tse (1992). An earnings prediction approach to examining intercompany information transfers. *Journal of Accounting and Economics* 15(4), 509–523.
- Gleason, C. A., N. T. Jenkins, and W. B. Johnson (2008). The contagion effects of accounting restatements. *The Accounting Review* 83(1), 83–110.
- Han, J. C. and J. J. Wild (1990). Unexpected earnings and intraindustry information transfers: Further evidence. *Journal of Accounting Research*, 211–219.
- Hoberg, G. and G. Phillips (2016). Text-based network industries and endogenous product differentiation. *Journal of Political Economy* 124(5), 1423–1465.
- Hrazdil, K. and T. Scott (2013). The role of industry classification in estimating discretionary accruals. *Review of Quantitative Finance and Accounting* 40(1), 15–39.
- Lang, L. H. and R. Stulz (1992). Contagion and competitive intra-industry effects of bankruptcy announcements: An empirical analysis. *Journal of financial economics* 32(1), 45–60.
- Lang, M. H. and R. J. Lundholm (1996). Corporate disclosure policy and analyst behavior. *The Accounting Review*, 467–492.
- Lee, C. M., P. Ma, and C. C. Wang (2015). Search-based peer firms: Aggregating investor perceptions through internet co-searches. *Journal of Financial Economics* 116(2), 410–431.
- O’Brien, P. C. and R. Bhushan (1990). Analyst following and institutional ownership. *Journal of Accounting Research*, 55–76.
- Olsen, C. and J. R. Dietrich (1985). Vertical information transfers: The association between retailers’ sales announcements and suppliers’ security returns. *Journal of Accounting Research*, 144–166.
- Petersen, M. A. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies* 22(1), 435–480.
- Pownall, G. and G. Waymire (1989). Voluntary disclosure choice and earnings information transfer. *Journal of Accounting Research*, 85–105.
- Pyo, Y. and S. Lustgarten (1990). Differential intra-industry information transfer associated with management earnings forecasts. *Journal of Accounting and Economics* 13(4), 365–379.
- Ramnath, S. (2002). Investor and analyst reactions to earnings announcements of related firms: An empirical analysis. *Journal of Accounting Research* 40(5), 1351–1376.

Veldkamp, L. L. (2006). Information markets and the comovement of asset prices. *Review of Economic Studies* 73(3), 823–845.

Figures

Figure 1

This figure plots our measure of closeness for five of the GIC6 industries over our sample time period. We ranked the industries on mean \mathcal{C} over our sample time period and randomly chose one GIC6 industry from each of the 5 resulting quintiles.

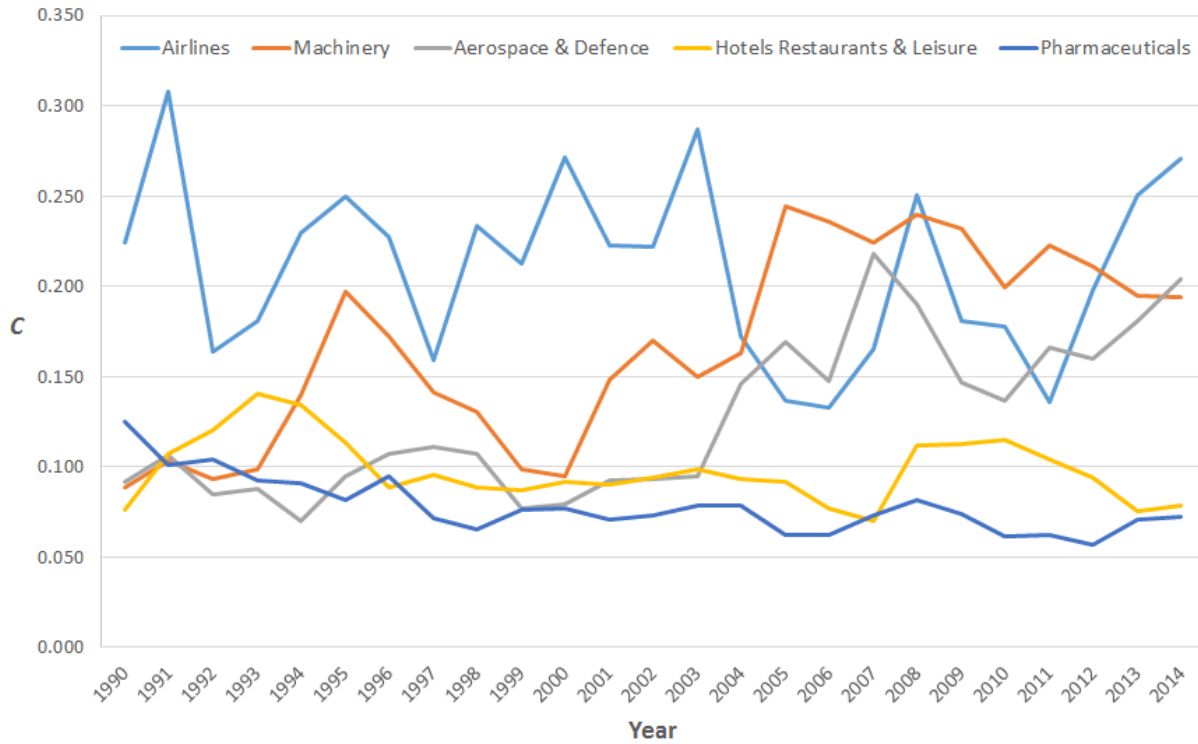


Figure 2

This figure plots the mean of our measure of closeness across all GIC6 industries for each year of our sample time period (1990-2014). For a given year, an industry's C is measured using the previous 12 quarterly earnings and cash flow numbers for each firm in that industry as of the end of the year following the measure description in the paper.

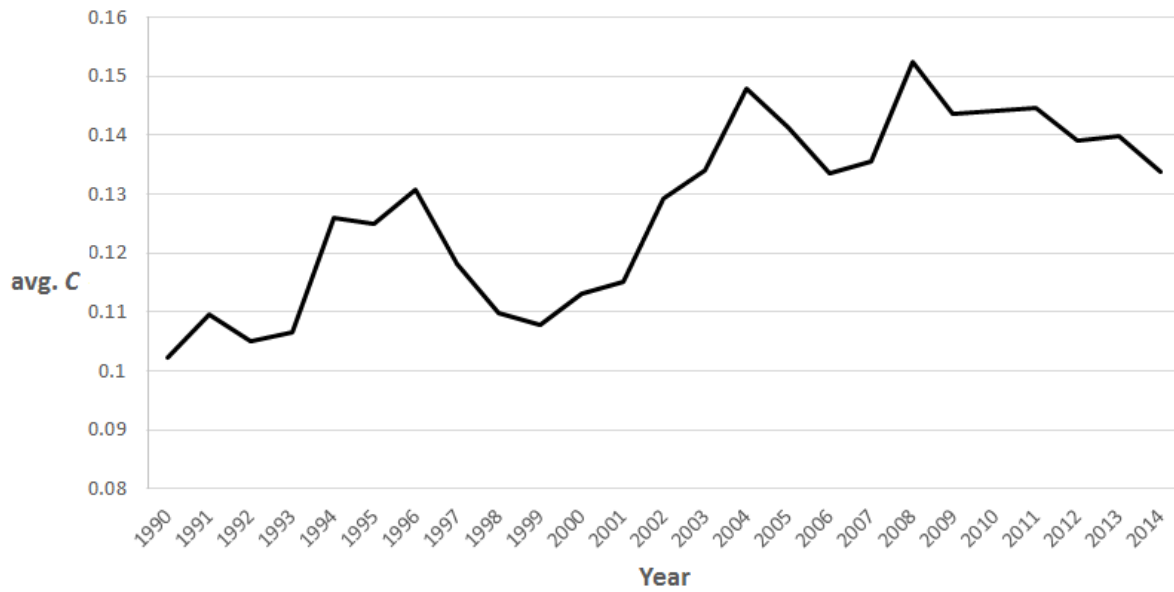
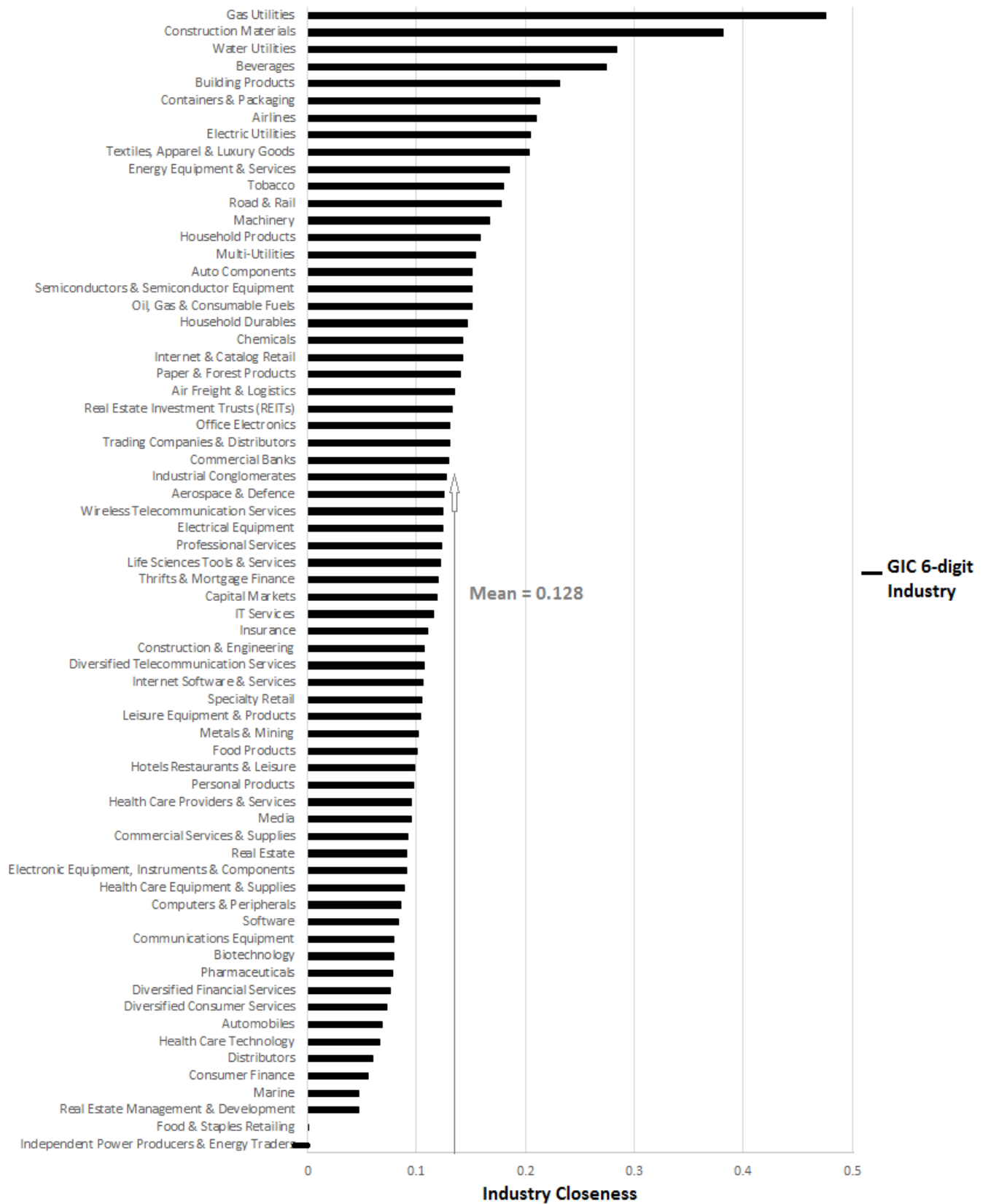


Figure 3

This figure plots the mean closeness of each GIC6 industry over our sample time period (1990-2014).



Tables

Table 1
Industrial Classification Schemes and Closeness

This table reports the pooled mean of our closeness measure, \mathcal{C} , across the industry-years as well as the pooled mean number of firms per industry-year for each of the schemes over our sample time period. This table also reports the number of industries represented as a fraction of the total number of industries per scheme. An industry group is not included if it has fewer than 5 firms. The sample time period is 1990-2014. The schemes are ranked in terms of \mathcal{C} and t-statistics from one-tailed hypothesis tests regarding the difference between the GIC6 average \mathcal{C} and the other schemes are reported below each scheme's mean \mathcal{C} . Significance at the 1%(5%)(10%) levels are signified by *(**)(***) respectively.

Scheme	avg. \mathcal{C}	avg. N	Representation
GIC6	0.128	34	67/67
FF	0.125 (0.15)	46	48/48
SIC2	0.121 (0.55)	34	62/83
SIC4	0.121 (0.56)	11	202/450
NAICS3	0.120 (1.01)	28	76/103

Table 2
Industrial Classification Schemes Versus Random Groupings

This table reports the pooled mean of our closeness measure, \mathcal{C} , across the industry-years and the pooled mean of \mathcal{C} for randomly chosen groups of similar size. Specifically, as of the end of each year of our sample time period we randomly select (with replacement) 1000 groups of firms of the same size as the average industry sizes of each of the schemes in our sample respectively and calculate the mean \mathcal{C} of the 1000 groups. This table reports these results averaged across the years. The sample time period is 1990-2014.

	Random	GIC6	FF	SIC2	SIC4	NAICS3
avg. \mathcal{C}	0.082	0.128 (13.07) ^{***}	0.125 (11.64) ^{***}	0.121 (10.50) ^{***}	0.121 (17.33) ^{***}	0.120 (10.53) ^{***}
avg. N	31	34	46	34	11	28

Table 3
Size Descriptive Statistics

This table reports pooled descriptive statistics for our sample versus the COMPUSTAT population (in parentheses) of firms with annual earnings and asset data over the same time period. The sample time period is 1990-2014. In computing the descriptive statistics we first delete the extreme 2 percentiles of the distribution. Data is in \$ millions

Variable	No. of Obs.	Mean	STD	10th%	Median	90th%
Earnings	81,234 (232,391)	\$36.77 (\$45.05)	\$124.01 (\$155.24)	-\$22.75 (-\$16.13)	\$3.20 (\$1.63)	\$121.05 (\$127.29)
Assets	94,083 (259,912)	\$1,982 (\$2,175)	\$46,423 (\$62,244)	\$18.34 (\$6.33)	\$355.25 (\$218.25)	\$5,069 (\$5,004)

Table 4
Within-Industry Contagion Effects from Leader Earnings Announcements

This table reports the mean(median) cumulative abnormal returns (CARs) of peer firms for one, two and three day windows following their industry leader's earnings announcement date. The CARs are classified according to whether the leader reported good news or bad news. The CARs of the leaders proxy for "good" and "bad" news. The mean(median) CARs of the leaders are also reported for the good and bad news samples. The sample time period is 1990-2014. T-statistics from a test of whether the mean CARs are significantly different from zero are reported below their respective mean CARs and ***(**)(*) represents significance at the 1%(5%)(10%) level respectively.

	Peer N	Leader N	Mean News	Window	Mean CAR	Med. CAR
Good News	18,259	731	2.53% (21.0)***	[0,1]	0.28% (8.36)***	0.05%
				[0,2]	0.33% (7.77)***	0.00%
				[0,3]	0.22% (4.40)***	0.10%
Bad News	17,065	690	-2.38% (-24.3)***	[0,1]	-0.21% (-6.30)***	-0.28%
				[0,2]	-0.1% (-2.26)**	-0.29%
				[0,3]	-0.05% (-0.88)	-0.40%

Table 5
Industry Contagion Effects Within Extreme Closeness Deciles

This table examines the contagion effects of industry leader earnings announcements for the extreme industry closeness deciles (D1 & D10). Panel A summarizes the results from the good news sub-sample and Panel B summarizes results from the bad news subsample. T-statistics from a test of whether the mean CARs are significantly different from zero for the mean news of the leaders and for the mean CARs of the peers are reported below their respective means. The sample time period is 1990-2014. T-statistics from a test of whether the difference between closeness decile mean CARs for the peer firms is statistically different from zero are reported below their respective differences and ***(**)(*) represents significance at the 1%(5%)(10%) level respectively.

Panel A: “Good News” Contagion Effects Within Extreme Closeness Deciles

	Closeness Decile 1	Closeness Decile 10	Difference
Closeness (\mathcal{C})	≤ 0.0699	≥ 0.2217	–
Leader N	47	50	–
Mean Leader News	0.43% (2.38)***	0.11% (0.68)	–
Peer N	1,043	1,067	–
Peer Mean CAR: [0,1]	0.04% (0.68)	0.75% (2.96)***	0.69% (2.08)**
Peer Mean CAR: [0,2]	0.02% (0.31)	0.46% (2.10)**	0.44% (1.91)**
Peer Mean CAR: [0,3]	0.03% (0.52)	0.65% (2.74)**	0.62% (2.01)**

Panel B: “Bad News” Contagion Effects Within Extreme Closeness Deciles

	Closeness Decile 1	Closeness Decile 10	Difference
Closeness (\mathcal{C})	≤ 0.0699	≥ 0.2217	–
Leader N	42	39	–
Mean Leader News	-0.27% (0.78)	-1.13% (-5.37)***	–
Peer N	1,402	673	–
Peer Mean CAR: [0,1]	-0.01% (-0.28)	-1.08% (-3.76)***	1.07% (3.42)***
Peer Mean CAR: [0,2]	-0.03% (-0.51)	-0.98% (-2.60)***	0.95% (-2.32)**
Peer Mean CAR: [0,3]	-0.06% (-0.78)	-0.91% (-2.23)**	0.85% (-2.21)**

Table 6
Interaction of Earnings Announcement Contagion Effect and Closeness

This table reports the coefficient estimates and adjusted R^2 values of a pooled estimation of regression equation (3) over three windows following leader earnings announcements. The dependent variable is the cumulative abnormal return of peer firm i over the window subsequent to the leader's earnings announcement. Panel A reports the results where we measure closeness using our new measure, C from equation (2). Panel B reports the results where we measure closeness using $1 - C_2$ and Panel C reports the results where we measure closeness using C_1 (the measure from previous literature). The sample time period is 1990-2014 and standard errors are clustered at the firm and year level. N is the number of regression observations. T-statistics from a test of whether the coefficient estimates are statistically different from zero are reported below their respective coefficients. ***(**)(*) represents significance at the 1%(5%)(10%) level respectively.

Panel A: Using C						
Window	Intercept	C	News	$C * \text{News}$	adj(R^2)	N
[0, 1]	-0.000 (-0.04)	0.003 (0.51)	-0.058 (-2.18)**	0.939 (5.41)***	0.54%	35,324
[0, 2]	-0.000 (-0.32)	0.011 (1.30)	-0.143 (-4.09)***	1.583 (6.98)***	0.49%	35,324
[0, 3]	0.000 (0.19)	0.004 (0.38)	-0.240 (-5.70)***	2.176 (7.99)***	0.36%	35,324

Panel B: Using $1 - C_2$						
Window	Intercept	$1 - C_2$	News	$(1 - C_2) * \text{News}$	adj(R^2)	N
[0, 1]	0.000 (0.66)	0.000 (0.22)	-0.002 (-0.13)	0.702 (4.75)***	0.50%	35,324
[0, 2]	0.000 (0.34)	0.004 (0.66)	-0.038 (-1.94)*	1.10 (5.42)***	0.47%	35,324
[0, 3]	0.001 (1.28)	-0.002 (-0.25)	-0.063 (-2.70)***	1.242 (6.48)***	0.33%	35,324

Panel C: Using C_1						
Window	Intercept	C_1	News	$C_1 * \text{News}$	adj(R^2)	N
[0, 1]	0.000 (0.67)	0.000 (0.13)	-0.032 (-1.91)*	0.543 (3.68)***	0.49%	35,324
[0, 2]	0.000 (0.28)	0.002 (0.37)	-0.020 (-1.48)	0.793 (4.52)***	0.45%	35,324
[0, 3]	0.000 (0.84)	-0.001 (-0.24)	-0.006 (-0.40)	0.852 (5.05)***	0.30%	35,324

Table 7
GL (2008) Replication Results

This table reports results from our replication of Gleason et al. (2008) along with the results from Table 1 Panel A of GL (2008) in parentheses for comparison. *** indicates a statistically significant two-tailed t-test of the null hypothesis that the mean cumulative abnormal return is not different from zero at the 1% level.

	<u>Restatement Firms (CARs)</u>			<u>Peer Firms (CARs)</u>		
	<u>n</u>	<u>Mean</u>	<u>Median</u>	<u>n</u>	<u>Mean</u>	<u>Median</u>
Panel A: Restatement Ret.						
Announcement [-1,+1]	161	-15.9%	-12.0%	3,251	-2.8%***	-2.0%
	(380)	(-19.8%)	(-14.6%)	(22,510)	(-0.5%***)	(-0.5%)
Pre-announce [-10,-2]	161	-1.1%	-2.0%	3,251	-1.1%***	-2.5%
	(380)	(-4.6%)	(-3.8%)	(22,510)	(-0.8%***)	(-0.8%)
Post-announce [+2,+10]	161	-2.2%	-1.7%	3,251	-1.5%***	-2.5%
	(380)	(-2.1%)	(-0.7%)	(22,510)	(-0.7%***)	(-0.7%)

Table 8
Interaction of Accounting Restatement Contagion Effect and Closeness

This table reports the coefficient estimates and adjusted R^2 values of a pooled estimation of regression equation (3) over three windows following accounting restatements identified by the Government Accountability Office (GAO) over the period 1997-2002. The dependent variable is the cumulative abnormal return of peer firm i over the window subsequent to the accounting restatement announcement. Standard errors are clustered at the firm and year level. N is the number of regression observations. T-statistics from a test of whether the coefficient estimates are statistically different from zero are reported below their respective coefficients. ***(**) represents significance at the 1%(5%) level respectively.

Window	Intercept	\mathcal{C}	News	$\mathcal{C} * \text{News}$	adj(R^2)	N
[0, 1]	-0.000 (-0.10)	-0.048 (-1.15)	-0.192 (-2.00)**	1.603 (2.24)**	0.44%	3,251
[0, 2]	-0.040 (-4.73)***	0.150 (2.67)***	-1.163 (-8.89)***	8.783 (9.048)***	2.81%	3,251
[0, 3]	-0.042 (-4.35)***	0.189 (2.93)***	-1.209 (-8.08)***	9.002 (8.11)***	2.14%	3,251

Table 9
Analyst Coverage and Industry Closeness

This table summarizes the results from regression equation (4) estimated as a panel where the standard errors are clustered at the firm and year level. The dependent variable is the natural logarithm of the number of analysts issuing an earnings per share forecast for firm i for period t using analyst's earnings forecast data from I/B/E/S. The sample time period is 1990-2014. T-statistics from one-tailed hypothesis tests are below their respective coefficients. ***(*) indicate significance at the 1%(10%) level respectively.

Variable	Prediction	Coefficients
<i>C</i>	+	0.801 (8.56)***
<i>Size</i>	+	0.268 (72.60)***
<i>Book-to-Market</i>	-	0.003 (1.07)
<i>Volume</i>	+	0.179 (50.22)***
<i>R&D</i>	+	0.000 (0.49)
<i>Depreciation</i>	+	-0.001 (-0.77)
<i>Issue</i>	+	0.052 (1.65)*
<i>Predictability</i>	+	0.138 (9.51)***
<i>Volatility Earn.</i>	-	-0.002 (-15.11)***
<i>Volatility Ret.</i>	-	-0.766 (-12.75)***
Adj. R^2		62.3%
Firm-year obs.		28,313

Table 10
Variable Definitions

This table gives the definition of all variables used in our study.

Variable	Definition
$Book\text{-}to\text{-}Market_{it-1}$	Ratio of the book value to the market value of equity for firm i measured at the end of year $t - 1$.
C_1	The mean of all pair-wise correlations in earnings and cash flows for a group of firms over a given time period (12 previous quarters in our analyses).
C_2	The mean of all pair-wise earnings and cash flow correlation matrix determinants for a group of firms over a given time period (12 previous quarters in our analyses).
C_j	The mean of C_1 and $1 - C_2$ if C_1 is positive and the mean of C_1 and $-(1 - C_2)$ if C_1 is negative for industry j using the earnings and cash flows for each firm over a given time period (12 previous quarters in our analyses). Earnings are variable IBQ in COMPUSTAT and cash flows are calculated by taking the change in COMPUSTAT variable $OANCFY$.
CAR_{iw}	Cumulative abnormal return for peer firm i over window $w \in \{[0, 1], [0, 2], [0, 3]\}$ around leader's(other firm's) earnings announcement(accounting restatement) date ($w = [0, 1]$).
$Coverage_{it}$	The natural logarithm of the number of analysts issuing an earnings per share forecast for firm i for period t .
$Depreciation_{it-1}$	Firm i 's depreciation expense in period $t - 1$ scaled by total sales, less the respective 6-digit GIC industry mean value of depreciation expense scaled by total sales.
$Issue_{it-1}$	Indicator variable that equals one if firm i issues debt or equity securities during the preceding, current or following year, zero otherwise.
$News_j$	Cumulative abnormal return for the leader(firm) of industry j on their earnings announcement(accounting restatement) date.
$Predictability_{it-1}$	R^2 of a regression of firm i quarterly earnings on firm i prior-year quarterly earnings using earnings data over the previous 12 quarters from $(t-3)$ – $(t-1)$.
R_{iw}	Cumulative return for peer firm i over window $w \in \{[0, 1], [0, 2], [0, 3]\}$ around leader's(other firm's) earnings announcement(accounting restatement) date ($w = [0, 1]$).
R_{Mw}	Cumulative value-weighted (NYSE/NASDAQ/AMEX) market return over window $w \in \{[0, 1], [0, 2], [0, 3]\}$ around leader's(other firm's) earnings announcement(accounting restatement) date ($w = [0, 1]$).
$R\&D_{it-1}$	Firm i 's research and development expense in period $t - 1$ scaled by total sales, less the respective 6-digit GIC industry mean value of research and development expense scaled by total sales.
$Size_{it-1}$	Natural logarithm of the market value of equity for firm i measured at the end of year $t - 1$.
$Volatility\ Earn_{it-1}$	Standard deviation of the 12 quarterly earnings using years $(t - 3)$ – $(t - 1)$ data for firm i .
$Volatility\ Ret_{it-1}$	Standard deviation of 36 monthly returns using years $(t - 3)$ – $(t - 1)$ data for firm i .