1-1-2017

Comovement, Financial Reporting Complexity, and Information Markets: Evidence from the Effect of Changes in 10-Q Lengths on Internet Search Volumes and Peer Correlations

Joshua J. Filzen
Boise State University

Maria Gabriela Schutte
University of Dayton

Comovement, financial reporting complexity, and information markets:

Evidence from the effect of changes in 10-Q lengths on internet search volumes and peer correlations

Joshua J. Filzen[*]
College of Business and Economics
Boise State University
1910 University Drive
Boise, ID 83725-1610
joshuafilzen@boisestate.edu

Maria Gabriela Schutte
Department of Economics and Finance
University of Dayton
Miriam Hall 502
300 College Park
Dayton, OH 45469
mschutte2@udayton.edu

Abstract: We investigate the effect of financial reporting complexity on stock comovement. We hypothesize that investors deal with complexity increases by acquiring low cost information. This information is typically informative not just about the firm of interest but also about other firms with similar fundamentals, which generates excess comovement. We find that increases in 10-Q word counts, a complexity proxy, are consistently followed by increases in 1) internet searches about the firm and 2) R^2s from regressions between the firm's returns and its peers'. On a large scale, complexity-induced comovement might hinder investors' ability to discriminate across stocks and identify business innovators.

Keywords: financial reporting complexity, comovement, information processing costs, complementarities, information acquisition.

JEL Classification: M41, D83, G12

* Corresponding author: Ph: (208) 426-3423.

"The topic of disclosure overload and complexity has been addressed by multiple organizations over many years. Some users of financial information seem to have an insatiable appetite for more information. Others observe that finding the truly significant information among the volume of routine and otherwise uninformative information is a challenge."

Disclosure Overload and Complexity: Hidden in Plain Sight
KPMG, 2011

1. Introduction

The conciseness and readability of financial reports has long been a topic of discussion among preparers, regulators, and users of financial information (FASB, 2008; SEC, 2008; KPMG, 2011; PWC 2011). Globalization, financial innovation, and an ever increasing number of accounting rules and disclosure standards have given rise to more complex financial reports. Although new standards such as those in the 2002 Sarbanes Oxley Act (SOX) have made financial reports more transparent, they have also made them more complex, thus hindering investors' ability to extract relevant information.

The objective of this study is to examine how unexpected increases in financial reporting complexity (or complexity shocks) affect stock return comovement, an important price informativeness proxy. So far, the literature has looked at complexity as a market friction that costs investors time and effort (Bloomfield, 2002; Li, 2008; Miller, 2010; Lee 2012). Because of its toll on information processing, several studies suggest that complexity shocks reduce price adjustment speeds (Lee, 2012; Cohen and Lou, 2012; You and Zhang, 2009). By delaying price adjustment, complexity shocks should increase idiosyncratic noise, reduce comovement, and make prices less informative. We propose that while this might be so, complexity might also affect comovement through a different channel. Because of its endogenous relation with information markets, complexity shocks might increase the demand for the private signals contained in low priced information, which is typically informative about multiple assets, thus

2

causing comovement to increase. The influence of complexity on comovement should be the net of these opposing effects.

Traditional models have thought information markets to be exogenous to asset pricing. However, new theoretical work (e.g., Veldkamp, 2006; Chamley, 2007; Ganguli and Yang, 2009; Manzano and Vives, 2011; Mele and Sangiorgi, 2015) has proposed an interplay between information and capital markets which could help explain the full effect of complexity on comovement. In these new models investors adjust to changes in their information environment through increases or reductions in their demand for information. In this way investors help determine the equilibrium price and informativeness of the signals they later use to make investment decisions. Alti, Kaniel and Yoeli (2012) provide evidence of this interplay. They show that investors respond to the uncertainty on the precision of an information signal by waiting for other confirming news before establishing stock positions.

Given the relative inflexibility of information supply, which is determined by the size of the financial sector, short-run changes in the price and informativeness of signals are mostly demand-driven. A signal is more or less valuable (i.e. useful to identify mispricings) depending on the privacy of its content. Private information on a signal depends on the number of investors that gain access to it. High priced signals are only accessible to a few investors without budget constraints, which makes their content private and more useful, but out of the reach of most investors for budgetary reasons. Alternatively, low priced signals are accessible to many investors, which automatically make them less private and informative, but still somewhat useful.

Because complexity hinders an investor's ability to extract relevant information from a financial report, a complexity shock increases the value of private information from other

3

sources. Therefore, the rational response to a complexity shock should be the same response that

a sudden drop in information production would elicit: an increase in the demand for low-priced

signals by investors with a budget constraint, despite their limited private information content

(Veldkamp, 2006).[1] Examples of such signals are news releases issued by the firm (Barber and

Odean, 2008 and Yuan, 2015), sector analyses, and analyst reports (De Franco, Hope, and

Laroque, 2015; Muslu, Rebello, and Xu, 2014). We use Google search volumes to proxy for the

demand of this type of information (Da, Engelberg, and Gao, 2011) and the number of words in

the 10-Q report to proxy for complexity[2] (Lee, 2012). Consistent with Veldkamp's prediction we

find a positive relation between changes in 10-Q word counts and changes in the number of

Google searches around the week of the 10-Q filing (i.e. searches after the filing minus searches

before the filing), even after accounting for a positive trend in search volumes. The larger the

magnitude of the complexity shock, the larger the increase in internet searches after the filing.

The high demand for low-priced signals is a consequence of their value to many investors, as

these can be used to price not just one, but multiple assets[3]. According to Veldkamp (2006), an

increase in the demand for low-priced signals in information markets has an effect on capital

---

[1] Veldkamp (2006) addresses the fact that her model appears contradictory to the idea that investors might gravitate towards high priced signals that other investors are not purchasing. However, she shows that in an information market where investors have budgetary constraints, investors will naturally gravitate towards low priced signals useful for pricing multiple assets. Thus, our results depend on at least some investors having a budgetary constraint, which we believe is a realistic assumption.

[2] Although the Fog Index (a readability measure) has historically been the primary measure of financial reporting complexity, we rely on file length because it is consistent with recent literature (Lee, 2012; and Loughran and McDonald, 2014.) The Fog Index uses average sentence length and proportion of complex words as readability proxies. However, Loughran and McDonald (2014) point out that because of the particularities of business writing, the complex words component of the Fog Index does not appropriately capture readability in this setting. The authors compare document length to alternative measures to find that document length and size yield consistent results with considerably less measurement error.

[3] The relation between demand and equilibrium price of information signals is a consequence of the nature of information. Information is different from physical goods in two ways. First it is non-rival, which means that its use does not preclude others from also using it. Second, it has high fixed costs but it is cheap to replicate. As a result, information producers respond to demand surges by replicating the signal as many times as necessary to meet the demand while charging a lower per-unit price (Veldkamp, 2006, 2011).

markets by generating excess comovement. This happens because part of the return covariance on stocks does not come from having synchronized cash flows, but from being priced using the same signals. Therefore, a complexity shock should ultimately increase the covariance between the stock experiencing the shock and other stocks priced using the same low-price information. Results from our empirical tests support this conjecture.

Throughout our empirical analysis we center our attention on groups of peer stocks[4] as a way of identifying stocks priced with the same signals. Consistent with our hypothesis, increases in 10-Q word counts are followed by significant increases in the R^2s from the regression of a firm's returns on its four closest peers. The effect of the complexity shock on peer comovement is statistically and economically significant. All else being equal, a one-standard-deviation increase in 10-Q word count, results in 5% higher R^2s with peer returns in the three months following the report's release.

We designed our empirical tests with the purpose of isolating the effect of complexity shocks on comovement from changes caused by other contemporaneous changes in information supply, increases in cash flow correlations, and other market frictions[5]. In additional analysis, we test the robustness of our main results to potential selection bias driven by our peer identification method and alternative ways of measuring complexity shocks, and also find results that support our prediction.

---

[4] Peer stocks are those stocks from peer companies, or the companies used by sell-side analysts to determine the value of the stocks they cover. We identify peers following Bhojraj and Lee (2002) who rely on similar "warranted multiples". Warranted multiples are one-year-ahead predicted enterprise value to sales ratios conditional on various firm and industry measures.

[5] We use fixed-effects to account for omitted variables such as joint cash flow shocks (Barberis, Shleifer, and Wurgler, 2005), joint under reaction to firm-specific news (including earnings announcements as in Lee, 2012), joint changes in the complexity of the peers' business environment, changes in industry-specific accounting standards, and secular increases in the number of accounting and disclosure rules.

Although the results from our regressions suggest that complexity shocks are a source of excess comovement, this effect may be short lived, lasting only until the financial sector increases its production of private information to meet the higher demand. However, in a market with large and frequent complexity shocks, these could contribute to market-level inefficiencies and resource misallocation (Morck, Yeung, and Yu, 2000; Wurgler, 2000; Durnev, Morck, Yeung, and Zarowin, 2003; Durnev, Morck, and Yeung, 2004). For example, Hoberg and Phillips (2012) find that firms with unique products and business practices comove less with their peers. In this way, low peer correlations could be instrumental in identifying business innovators. In the presence of large and frequent complexity shocks, peer correlations might lose their efficacy as stock-picking tools.

Our study sheds new light on the interplay between capital and information markets. To the best of our knowledge we are the first to examine the role of complexity shocks on comovement, complementing what has already been done for opacity, another aspect of the firm's information environment (Jin and Myers, 2006; Hutton, Marcus, and Tehranian, 2009).[6] In the case of opacity (which is the opposite of transparency), comovement arises not because of surges in information demand but because of information asymmetries between managers and investors. With opaque information, investors see some of the firms' cash flow changes, but not all. This makes prices respond only to those firm-specific signals that can be seen, resulting in lower firm-specific volatility and higher comovement. Complexity is different from opacity as it does not impair market participants from seeing all signals. An information environment can be fully transparent yet complex if all signals are visible but investors cannot understand them. In such

---

[6] We note that a recent study by Datta, Iskandar-Datta, and Singh (2014) calls into question the robustness of the findings in Hutton et al. (2009), which furthers the need for additional work in this area.

cases investors will deal with complexity by acquiring low cost information. In the end, both complexity and opacity are sources of excess comovement.

2. How complexity shocks influence a firm's information environment

Information choice theory predicts that investors respond to complexity shocks by increasing their demand for low-price information, which ultimately increases excess comovement. However, a positive relation between current complexity changes and future comovement does not necessarily imply that investors are purchasing low-priced information in response to complexity shocks. Therefore, to truly test the theory, we need a direct test between changes in information complexity and subsequent changes in information demand.

There are many examples of low-priced signals that we could use for this analysis such as news releases issued by the firm, news and headlines from the popular press (Barber and Odean, 2008 and Yuan, 2015), sector analyses, and sell-side analyst reports (De Franco et al., 2015; Muslu et al., 2014). However, the fact that these reports are being produced by the financial sector does not guarantee that they are being requested and used by investors. Instead of relying on one or a combination of these signals, we use aggregate search frequency from Google (Da et al., 2011) to directly capture low-price information demand. We choose this measure for three reasons. First, investors nowadays use a search engine to look for information signals, and Google continues to be the most frequently used search engine. According to Net Marketshare, as of July of 2014, Google was the world's most popular search engine, with 68.75% of all searches worldwide[7]. Second, search is a revealed attention measure, what stocks investors are searching for in Google are without a doubt the stocks investors are demanding information for.

---

[7] Source: Net Marketshare (http://www.digitaltrends.com/web/google-baidu-are-the-worlds-most-popular-search-engines/ last accessed on November 23rd, 2015.)

Third, it is very inexpensive to use a search engine to find information signals and many of these signals are available at no cost. Hence, it is reasonable to assume that internet searches capture the demand for low-priced information.

We extract data from two Google products: Google Trends and Google AdWords and use it to build a sample of internet search volumes in weekly frequencies[8]. There are 32,753 observations in this sample with matching 10-Q filings available in machine-readable format covering the period 2004-2009. To test the relation between changes in financial statement complexity and changes in internet search volume, we examine the period 12 weeks before and 12 weeks after the 10-Q filing. Even though we use a relatively short window, we note that increases in search volume throughout time could influence our analysis. Therefore, we utilize the methodology in Baum (2006) to detrend the pre and post filing series' by estimating a regression with a count variable for time as the independent variable. In our tabulated analysis, we use the residuals from this estimation process to form a time independent series, and following Baum (2006) we add the mean of the series to the residuals to "rebench" the data for easier interpretation.[9]

Each quarter, we sort observations into quartiles according to changes in 10-Q word counts, our proxy for the magnitude of the complexity shock. The construction and statistical properties of this variable are described at length in section 3.3. Panel A in Table 1 presents the average of

_____

[8] Data to construct the Google search volume database was downloaded from Google Trends and Google AdWords during the first three weeks in May of 2013. The dataset contains search volumes on the Google Search engine for ticker symbols of all publicly traded companies in the NYSE, AMEX, and NASDAQ. We use the data extracted through these tools to produce a historical series of Google search volumes since the first week in January of 2004, the first month of data availability. For each ticker symbol in the WRDS monthly stock return file we extract the average Google monthly search volume for the past 12 months from Google Ad Words and convert it to a weekly frequency. We then determine the search volumes for the past 12 months. Google Trends does not give a historical times series of absolute search volumes, but an index where 100 indicates the week in which the search term reached its highest level. To convert the search volume indices to absolute numbers we calculate the average weekly search volume by multiplying the weekly index times the average weekly volume over the past 12 months. To calculate search volumes since January of 2004, which are on a different basis as those for the past 12 months, we use the conversion: (weekly index*Search volume in first week of most recent 12 month period)/Search index in first week of most recent 12 month period. Finally, we connect the search volumes since January 2004 with search volumes over the past 12 month period to obtain the complete series.
[9] We find qualitatively similar results using raw data, rather than de-trended data.

total internet searches (de-trended) on a firm's ticker symbol during the 12 weeks before and 12 weeks after the week of its 10-Q filing (excluding the filing week). We find an increase in the number of internet searches on a stock's ticker symbol during the 12 weeks following the week of a firm's 10-Q release for all levels of complexity shocks during our sample period, however the differences are statistically significant only in the two largest quartiles of changes in financial reporting complexity, consistent with our expectations. The change in search volumes ranges between 1,481 and 6,222 searches. Furthermore, consistent with our hypothesis, search volume changes following 10-Q releases are proportional to the size of the complexity shock. The larger the shock, the larger the change in the number of internet searches. Panel B in Table 1presents the average difference in mean weekly ticker searches in the 12 weeks before and after the filing. Consistent with Panel A, Panel B reports that the highest average change in search volume occurs in the highest quartile of changes in financial reporting complexity, and the lowest average change occurs in the lowest quartile. A t-test of the difference in this variable between the top and bottom quartiles shows this difference is statistically different from zero (*t*-stat =2.27; *p*-value=0.02.)

Table 1

Influence of Financial Reporting Complexity Shocks on Internet Search Volumes

Panel A presents the average of de-trended total internet search volume on a firm's ticker symbol during the 12 weeks before and 12 weeks after the week of its 10-Q filing. The total number of firm-quarters in the sample is 32,753 covering the period 2004-2009. The number of weekly internet searches is obtained from Google Trends and Google AdWords. Stocks are sorted into quartiles according to the change in 10-Q length, a proxy for financial reporting complexity. Panel B presents the average difference in mean weekly de-trended ticker searches in the 12 weeks around the 10-Q filing.

Panel A: Relation between change in 10-Q length and average de-trended total search volumes before and after the 10-Q release

		Change in 10Q Length (Quartile)			
	All Firms	Q_1	Q_2	Q_3	Q_4
Average internet searches before 10-Q release	515,381	534,466	491,388	511,391	524,278
Average internet searches after 10-Q release	518,800	535,947	493,163	515,590	530,500
Difference	3,419	1,481	1,775	4,199	6,222
Percent change	0.66%	0.28%	0.36%	0.82%	1.19%
p-value for difference	0.00	0.38	0.28	0.03	0.00

Panel B: Relation between change in 10-Q length and average change in mean weekly de-trended search volume

	All Firms	Change in 10Q Length (Quartile)				$\overline{Q4 - Q1}$	$t_{(Q4-Q1)}$	$P_{Ho:Q4-Q1=0}$
		Q_1	Q_2	Q_3	Q_4			
Avg. change in mean weekly de-trended internet searches	168	30.7	218.3	131.9	290.4	259.7	2.27	0.02

Overall, the numbers presented in Table 1 are consistent with an increase in low-price information demand following a complexity shock and with an information-choice explanation for the positive relation between complexity shocks and comovement. This relation will be tested and measured statistically in subsequent sections.

3. Measuring the influence of complexity shocks on comovement

3.1 Sample construction

The starting point of this sample is the intersection of Compustat and CRSP with 10-Q filings available in machine-readable format.[10] Analyst forecast data are retrieved from IBES, under the assumption that missing IBES data implies no analyst coverage for that period. Data on institutional ownership are obtained from the SEC's 13f reports available in the Thomson Financial's Institutional Holdings data set. Table 2 outlines the sample construction procedures. We begin by identifying peer groups following the methodology in Bhojraj and Lee (2002). Data requirements to perform the warranted multiple calculations are highly restrictive. From the 176,336 firm-years with available data in Compustat and CRSP, only 55,949 have the necessary data to calculate warranted multiples. Further, after limiting the sample to firms and peers with at least two consecutive 10-Qs with necessary stock return data (29,250 firm-quarters), and firms with available data for all treatment and control variables we obtain a final sample of 26,484 firm-quarters during the period 1995 through 2009.[11]

---

[10] We utilize the Python programming language to gather 10-Q documents and tabulate lexical properties. We first convert all filings to plain text, thus allowing us to include both HTML and plain text quarterly filings in our analysis.

[11] When assigning peer groups we do not impose restrictions on the potential peers based on data necessary in later analysis. This is to ensure we find the most valid peer group, and not just the set of peers conditional on data availability. We lose a substantial number of observations (53,942 firm-quarters) based on all members of a firm's peer group not having two consecutive 10-Qs in the period to calculate average peer group complexity, used as a control in our analysis. To make sure that our regression coefficients are representative of the universe and not

Table 2
Sample construction

	Observations Subtracted	Sub-total
Warranted Multiples Calculation		
Firm-years on Compustat between 1995-2009 with prior year data availability		176,336
Less: REITS and ADRs	(9,361)	166,975
Less: Firms with prices<$3	(40,232)	126,743
Less: Firms with negative equity	(7,610)	119,133
Less: Missing necessary data for warranted multiple calculation	(58,632)	60,501
Less: Trimming at 1 and 99% by year	(4,254)	56,247
Less: Industries with less than 5 members	(298)	55,949
Total Firm-years with Warranted Multiple	55,949	
Less: Firms with less than 10 potential peers	(5,398)	50,551
Total firm-years with 4 peers identified	50,551	
Total firm-quarters with 10-Q length, from the 50,551 firm-years	101,025	
Final Sample Selection		
Less: Firm-quarters missing necessary returns data surrounding the 10-Q filing for the firm and it's complete assigned peer group	(17,833)	83,192
Less: Firm-quarters without all four peers having 10-Q length data	(53,942)	29,250
Less: Firm-quarters without necessary data to construct control variables	(2,766)	26,484
Total number of firm-quarters used in analysis	26,484	

unique to the smaller sample, we relax the data requirements to calculate average peer complexity only for sets with at least 3 of the 4 peers having complexity data, as well as completely unrestricting that requirement. Results from these regressions are qualitatively the same as those presented.

Table 3 presents our sample's distribution across years and industry groups. For the sake of comparison, we present the distribution of the total number of firm-quarters in the CRSP/Compustat universe in a separate column. We observe that the distribution of observations in our sample across years and industries closely resemble the distribution in the CRSP/Compustat universe and do not detect substantial concentration of data points in any particular year or industry. Panel A shows the distribution of firm-quarters by fiscal year. The sample begins in 1995, which is the first year with significant machine-readable quarterly SEC filings. The smallest number of observations occurs on that first year of the series (82 firm-quarters, 0.31% of the sample) while the largest occurs in 1997 (3,016 firm-quarters, 11.39% of the sample). Panel B shows the sample distribution by industry. We sort firm-quarters according to the Fama and French 30 industry portfolios.[12] In general, the sample appears to have reasonable coverage across many industries. However, there is a slight overrepresentation towards firms in Personal and Business Services (4,319 firm quarters, 16.31% of the sample), Healthcare, Medical Equipment and Pharmaceuticals (3,931 firm-quarters, 14.84% of the sample), and Business Equipment (3,940 firm-quarters, 14.88% of the sample), and underrepresentation of Banking, Insurance, Real Estate and Trading (749 firm-quarters, 2.83% of the sample), and no representation in Tobacco Products and Precious Metals, Non-Metallic, and Metal Mining.

Table 3 also shows the average 10-Q word counts for firms in our sample broken down by year and by industry group. We observe a sustained increase in the number of words in the 10-Q statements over the years. At the onset of the sample period, in 1995 the average 10-Q had 8,184 words. At the end of the series, in 2009, this number had risen to 19,787 words. In other words,

---

[12] The 30 industry portfolios are available at Kenneth R. French's data library (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html, last accessed January 28, 2013).

the average 10-Q more than doubled its length in those 14 years. We use year fixed-effects to control for the positive trend in average 10-Q length in our regressions. We also find some distinguishable patterns in 10-Q word counts across industries. Two industries in particular, Coal and Utilities have abnormally long 10-Qs with 24,625 and 25,048 words, respectively. Industries with relatively short 10-Qs are Textiles (8,168 words), Automobiles and Trucks (9,568 words) and Apparel (9,650 words). We use industry fixed-effects in our regressions to control for the differences in 10-Q word counts across industries.

Table 3
Sample distribution
This table presents information about the distribution of 10-Q lengths across time and industry.
Industry classifications are the 30 industry groups in Fama and French (1997).

	Average 10-Q Word Count	Sample Distribution		CRSP/Compustat Distribution	
		Frequency	Percent	Frequency	Percent
Panel A: Sample Distribution by Fiscal Year					
1995	8,184	82	0.31	33,794	6.89
1996	9,646	1,288	4.86	35,767	7.29
1997	9,968	3,016	11.39	36,108	7.36
1998	11,092	2,815	10.63	36,031	7.35
1999	12,243	2,451	9.25	36,796	7.50
2000	9,252	1,618	6.11	36,614	7.47
2001	8,279	1,216	4.59	34,521	7.04
2002	11,233	1,601	6.05	32,685	6.67
2003	13,843	2,458	9.28	31,147	6.35
2004	14,432	1,552	5.86	32,501	6.63
2005	15,668	1,456	5.50	31,105	6.34
2006	18,847	1,476	5.57	29,903	6.10
2007	17,247	1,806	6.82	28,938	5.90
2008	18,987	2,479	9.36	27,733	5.66
2009	19,787	1,170	4.42	26,676	5.44
Total	13,247	26,484	100.0	490,319	100.0
Panel B: Sample Distribution by Industry					
Food Products	10,925	483	1.82	9,462	1.93
Beer & Liquor	10,417	112	0.42	1,528	0.31
Tobacco Products	N/A	0	0.00	512	0.10
Recreation	14,510	625	2.36	12,628	2.58
Printing and Publishing	11,790	440	1.66	4,954	1.01
Consumer Goods	10,983	468	1.77	6,993	1.43
Apparel	9,650	288	1.09	5,528	1.13
Healthcare, Medical Equipment, Pharmaceuticals	13,569	3,931	14.84	49,185	10.03
Chemicals	13,444	807	3.05	8,103	1.65
Textiles	8,168	48	0.18	1,910	0.39
Construction and Construction Materials	10,561	830	3.13	12,566	2.56
Steel Works Etc.	12,677	447	1.69	6,021	1.23
Fabricated Products and Machinery	10,666	1,327	5.01	15,401	3.14
Electrical Equipment	9,944	431	1.63	6,327	1.29
Automobiles and Trucks	9,568	407	1.54	6,159	1.26
Aircraft, ships, and railroad equipment	11,237	220	0.83	2,609	0.53
Precious Metals, Non-Metallic, and Metal Mining	N/A	0	0.00	10,867	2.22
Coal	24,625	16	0.06	748	0.15
Petroleum and Natural Gas	14,993	1,533	5.79	20,712	4.22
Utilities	25,048	684	2.58	15,826	3.23
Communication	18,416	577	2.18	18,047	3.68
Personal and Business Services	14,931	4,319	16.31	62,954	12.84
Business Equipment	13,284	3,940	14.88	55,685	11.36
Business Supplies and Shipping Containers	10,953	377	1.42	6,514	1.33
Transportation	11,579	764	2.88	11,758	2.40
Wholesale	11,986	1,179	4.45	16,677	3.40
Retail	12,758	637	2.41	20,501	4.18
Restaurants, Hotels, Motels	13,392	231	0.87	8,815	1.80
Banking, Insurance, Real Estate, Trading	13,212	749	2.83	76,179	15.54
Everything Else	11,743	614	2.32	15,150	3.09
Total	13,037	26,484	100.0	490,319	100.0

3.2 Peer group formation

We hypothesize that unanticipated shocks to financial reporting complexity make firm returns more strongly correlated with those of its peers. We identify a firm's peer group following the method in Bhojraj and Lee (2002), which identifies peers according to their predicted enterprise value to sales ratios, also known as their warranted multiples. Warranted multiples are calculated by fitting annual cross-sectional regressions of enterprise value to sales on eight explanatory variables.[13] Coefficients from these regressions are used to estimate the one-year-ahead predicted enterprise value to sales.[14] A firm's peer group consists of the four industry peers with the closest warranted multiples. Because this process uses annual data, we assume that the peer groups remain effective for all quarters during the fiscal year.

Although we follow Bhojraj and Lee (2002) as closely as we can, our experimental design requires that we make four modifications to their method:

1) We require that a CRSP permno be available for all stocks at the start of the sample selection process.[15]

2) Given that peer comparisons do not require that peer firms have an analyst following, we replace the analyst-generated growth forecasts in the Bhojraj and Lee (2002) model with realized sales growth over the previous year.

3) We relax the minimum sales requirement of $100 million. Although relaxing this restriction should help incorporate smaller firms into the sample, the sample remains somewhat skewed towards larger firms because of other extensive data requirements.

---

[13] These variables are: average price-to-book ratio for the firm's industry; industry-adjusted profit margin; a variable that captures the differential effect of profit margin on the price-to-sales ratio of loss firms; industry-adjusted growth forecasts; book-value leverage ratio; return on net operating assets; return on equity; and research and development to sales ratio. See Bhojraj and Lee (2002) for a detailed description of the estimation process.

[14] The coefficient estimates we obtain from regressing warranted multiples on their determinants are similar to those in Bhojraj and Lee (2002) (untabulated).

[15] We do not require that returns data be available at this stage to obtain the closest possible matches.

4) To ensure that peer groups are not merely a function of the remaining firms in a given industry, we require that at least 10 potential peers be available in each industry before selecting the four closest matches.

3.3 Measurement of financial reporting complexity shocks

We use quarterly changes in 10-Q lengths to proxy for the size of the complexity shock faced by economic agents when attempting to extract relevant information from a new 10-Q report (Lee, 2012; Brown and Tucker, 2011).[16,17] Prior research has primarily relied on two readability measures to proxy for financial reporting complexity: the length of SEC filings and the Gunning Fog Index (Li, 2008; You and Zhang, 2009; Miller, 2010; Lehavy, Li, and Merkley, 2011; Lee, 2012). Loughran and McDonald (2014) find that the Fog Index is poorly specified when used to evaluate financial texts, and recommend document length as an alternative. We utilize the number of words in the SEC 10-Q filings as our primary complexity measure to allow for comparability with Lee (2012).[18] We gather all 10-Q filings from Direct Edgar for the years 1995-2009 using the Python programming language, and count the number of words contained in each file. We believe that changes in 10-Q word counts (as opposed to levels) are a better proxy

---

[16] We utilize quarterly data in our analysis to maximize the power of our tests, because we believe this is the most likely reporting window where increases in length will be unexpected to analysts. Lee (2012) finds that the information processing costs associated with increases in 10-Q complexity are associated with market under-reaction. This is consistent with complexity affecting investors at quarterly and not just annual intervals, as expected. We examine changes in an annual reporting window in section 5.2.

[17] We test the robustness of our results to an alternative complexity shock definition. In particular we consider whether investors experience a shock only when the change in 10-Q length exceeds the change of other stocks in the industry. We redo our main analysis using the difference in industry-quarter mean adjusted 10-Q word counts as our primary independent variable of interest. Results from this additional test are similar to those presented in Table 6.

[18] In the final version of Loughran and McDonald (2014) the authors recommend using the 10-K complete submission file size as a measure for reporting complexity because it is exceptionally easy to determine, not necessarily due to superior performance relative to the number of words in the SEC submission document. In previous, unpublished versions of their study, the recommended proxy was the number of words in the SEC submission document. We report the results from word counts, so the economic interpretation of our results would remain consistent with that in Lee (2012). Results obtained using 10-Q file size as the complexity proxy instead of 10-Q word count are similar to those presented in Table 6.

of unanticipated shocks on financial reporting complexity, which is the treatment variable in our hypothesis.[19] Table 4 reports the average 10-Q in the sample has 13,436 words (10,487 for the median). 10-Q word counts range between 2,278 and 62,862 words. Average and median word counts for stocks in our sample are slightly lower, yet comparable, to those in Lee (2012), who report median 10-Q counts of 14,499 words. $COMPLEX_{it}$ is the natural logarithm of the number of words in the 10-Q. $\Delta COMPLEX_{it}$ is the change in $COMPLEX_{it}$ between previous and current quarters and is the measure of financial reporting complexity used throughout our analysis. The average quarterly 10-Q change in the sample is 452 (490 for the median) words (untabulated). Table 4 includes descriptive statistics on our variables of interest. Compared to the average $COMPLEX_{it}$ of 9.253, the average mean $\Delta COMPLEX_{it}$ of 0.044 seems very small and suggests stickiness in 10-Q filings from one quarter to the next. However, the standard deviation of 0.57, which corresponds to a standard deviation in 10-Q changes of 10,061 words (untabulated), indicates substantial variation in this variable.

---

[19] Because 10-Qs are only produced in the first through third quarters of the fiscal year we measure the first quarter change as the first quarter length minus the third quarter length in the prior year.

Table 4

Sample statistics for variables used to determine the influence of complexity on peer stock comovement

This table reports sample statistics for the variables used in our analysis. We report number of observations (N), mean, standard deviation, and key points in the distribution. ΔSYNCH$_{it}$ is the change in peer group synchronicity around the filing date of the 10-Q, where synchronicity is measured as a log transformation of the R^2 from a regression of firm returns on value-weighted peer group returns. ΔCOMPLEX$_{it}$ is the change from last quarter of the natural logarithm of the number of words in the 10-Q report. All other variables are described in detail in the Appendix.

Variable	N	Mean	Standard Dev.	Min	p25	p50	p75	Max
R^2	26,484	0.106	0.143	0.000	0.009	0.045	0.144	0.642
$SYNCH_{it}$	26,484	-3.420	2.385	-11.320	-4.680	-3.063	-1.738	0.627
$\Delta SYNCH_{it}$	26,484	0.028	2.526	-7.547	-1.228	0.031	1.312	7.475
RAW_WORDS_{it}	26,484	13,436	10,487	2,278	6,073	10,600	17,452	62,862
$COMPLEX_{it}$	26,484	9.253	0.711	7.731	8.712	9.269	9.767	11.050
$\Delta COMPLEX_{it}$	26,484	0.044	0.565	-2.168	-0.087	0.054	0.186	2.255
$AVG\_PR\_COMPLEX_{it}$	26,484	9.249	0.510	7.967	8.872	9.288	9.638	10.360
$\Delta AVG\_PR\_COMPLEX_{it}$	26,484	0.043	0.323	-1.191	-0.118	0.055	0.192	1.278
$QSIZE_{it}$	26,484	5.942	1.866	2.147	4.581	5.837	7.178	10.860
$\Delta QSIZE_{it}$	26,484	0.016	0.317	-0.958	-0.145	0.020	0.184	0.978
MKT_SHARE_{it}	26,484	0.074	0.157	0.000	0.002	0.012	0.062	0.896
ΔMKT_SHARE_{it}	26,484	0.000	0.019	-0.093	-0.001	0.000	0.001	0.103
$ANALYSTS_{it}$	26,484	1.395	0.946	0.000	0.693	1.386	2.079	3.258
$\Delta ANALYSTS_{it}$	26,484	0.028	0.253	-0.693	0.000	0.000	0.095	1.099
$INSTPERC_{it}$	26,484	0.330	0.314	0.000	0.000	0.265	0.590	1.000
$\Delta INSTPERC_{it}$	26,484	0.010	0.058	-0.154	-0.004	0.000	0.019	0.299
$QBTM_{it}$	26,484	-0.874	0.732	-2.921	-1.322	-0.820	-0.370	0.735
$\Delta QBTM_{it}$	26,484	0.024	0.323	-0.918	-0.146	0.013	0.182	1.094
ROE_{it}	26,484	0.014	0.071	-0.334	0.002	0.024	0.044	0.208
ΔROE_{it}	26,484	-0.001	0.069	-0.343	-0.014	0.000	0.012	0.342
$LEVERAGE_{it}$	26,484	0.438	0.211	0.054	0.262	0.442	0.603	0.895
$\Delta LEVERAGE_{it}$	26,484	0.001	0.052	-0.249	-0.018	-0.002	0.017	0.209
DED_IOERC_{it}	26,484	0.042	0.076	0.000	0.000	0.000	0.058	0.339
ΔDED_IOPERC_{it}	26,484	-0.001	0.030	-0.152	0.000	0.000	0.000	0.112
$INFO_{it}$	26,484	0.165	0.372	0.000	0.000	0.000	0.000	1.000
$OPACITY_{it}$	26,484	0.002	0.039	-0.125	-0.017	0.001	0.020	0.138
$\Delta OPACITY_{it}$	26,484	0.039	0.043	0.000	0.011	0.024	0.050	0.240
$NOISE_{it}$	26,484	0.058	0.167	-0.320	-0.057	0.051	0.165	0.489
$\Delta NOISE_{it}$	26,484	-0.001	0.198	-0.480	-0.132	0.000	0.132	0.474
$TURNOVER_{it}$	26,484	1.528	1.627	0.033	0.443	0.999	1.998	8.953
$\Delta TURNOVER_{it}$	26,484	0.029	1.174	-4.295	-0.330	0.001	0.352	4.750

ΔAVG_PR_COMPLEX$_{it}$ is the quarterly change in *AVG_PR_COMPLEX$_{it}$*, or the average 10-Q word counts of the firm's four closest peers. Summary statistics of average peer 10-Q word counts are almost identical to those for *COMPLEX$_{it}$* and *ΔCOMPLEX$_{it}$*. This is to be expected given that the firms under study and their matching peers are drawn from the same sample. We include *ΔAVG_PR_COMPLEX$_{it}$* in our regressions to account for the changes in R^2 caused by changes in peer complexity. Consistent with information choice theory, increases in peer 10-Q word counts should increase the demand for low-cost information, resulting in a positive relation between changes in average peer 10-Q word counts and changes in peer stock comovement.

3.4 Measurement of peer group comovement

The dependent variable in our tests is the change in the correlation between a stock's returns and those of its four closest peers before and after the 10-Q filing. Based on our hypothesis, we expect that word count increases will be followed by increases in return comovement among peer stocks. Following prior literature (Morck et al., 2000; Piotroski and Roulstone, 2004; Chan and Hameed, 2006; and Crawford, Roulstone, and So, 2012; among others), we denote peer comovement as *SYNCH$_{it}$*, which is short for synchronicity.[20] We measure peer comovement changes by the difference between the R^2 of a stock's returns with the value-weighted average return of its four closest peers during the three months prior to the 10-Q filing and the three months following the 10-Q filing, excluding the month of the filing in both periods. Our measure is similar to Crawford et al. (2012); however, instead of looking at the R^2 of the

---

[20] The accounting and finance literatures have used several terms to describe the movement of stock prices in the same direction. The most commonly used terms are synchronicity, comovement, firm-specific return variation (to denote the portion of returns that do not co-vary), or simply R^2 (e.g., Roll, 1988; Jin and Myers, 2006). The term synchronicity is a relatively newer term, it was first used in Morck et al. (2000) to identify co-variation measured as the proportion of stock prices moving in the same direction in a given country.

regression of stock returns on the market portfolio's, we look at the R^2 of the regression of stock

returns on its four closest peers', as shown in equation (1):

$$RETURN_{it} = \gamma_0 + \gamma_1 PEER\_RETURN_{it} + \varepsilon_{it} \tag{1}$$

$RETURN_{it}$ is the daily return of firm i on day t. $PEER\_RETURN_{it}$ is the value-weighted average

daily return of the firm's four closest peers on day t. We fit this regression on each period before

and after the 10-Q filing with a minimum of 50 daily observations. The R^2 from this regression

captures the amount of peer information incorporated into the stock's returns and the extent to

which stock prices move together. The R^2 for the average firm-quarter is 0.11 (median of 0.045)

and the measure ranges between near zero (1.25×10^{-5}) and 0.64. We apply a log-transformation

to limit the influence of outliers and unbind the statistic, creating a closer approximation to the

normal distribution. Specifically, we use the transformation:

$$SYNCH_{it} = ln(R^2/(1-R^2)) \tag{2}$$

R^2 is the coefficient of determination from equation (1). Intuitively, $SYNCH_{it}$ measures the

degree to which firm-level stock returns are explained by peer group returns. Since we estimate

$SYNCH_{it}$ for the three months before and the three months after the filing date, $\Delta SYNCH_{it}$ is the

change in peer group synchronicity around the 10-Q filing.

3.5 Control variables

To isolate the effect of complexity shocks on peer comovement we introduce control

variables that proxy for five factors that could influence peer comovement at the time of the

shock:

1) Changes in information production.

2) Changes in transaction costs.

3) Joint cash flow shocks caused by simultaneous shocks to firm fundamentals. Examples of these shocks could be higher industrywide sales caused by aggregate income shocks, lower barriers to trade, increases in foreign demand, or the launching of new products.

4) Joint underreaction to firm-specific news, including earnings announcements (Lee 2012).

5) Trends in information supply, transaction costs, and firm fundamentals not captured by our control variables.

3.5.1 *Information Production*

Information production can be defined as all the news generated by the financial sector about a firm. In general, comovement declines with the production of firm-specific information and increases with the production of market and industry information. Firm-specific information reduces comovement because with more firm-specific information, the incentive to demand low-cost signals, containing mostly industry information, declines.

Two separate streams of literature produce opposing predictions on how (firm-specific) information production affects comovement. On the one hand, information choice theory, and Veldkamp (2006) in particular, predict that an increase in the number of information signals about a firm should reduce comovement. This is because a firm-specific news story, say for example a press release, is assumed to contain mostly firm-specific information and an increase in firm-specific information is expected to reduce investors' incentives to acquire low-cost signals. With fewer investors relying on low-cost signals, comovement should decline. On the other hand, the information transfers literature assumes that a firm-specific news story has a considerable amount of market and industry information embedded, which can be "transferred" to other similar firms (Foster, 1981; Clinch and Sinclair, 1987; Han and Wild, 1990; Olsen and Dietrich, 1985; Baginski, 1987; and Gleason, Jenkins, and Johnson, 2008; among many others).

23

For example, a positive earnings announcement for a retail firm is typically well received by all investors holding long positions in the retail sector because sales revenue across many retailers are highly positively correlated. Therefore, an increase in the production of firm-specific news increases comovement because with more news being generated, more information is transferred to other firms.

We include firm characteristics such as firm size, book-to-market ratio, and leverage as information production proxies (Hutton et al., 2009). Traditionally, information supply has been captured by firm size, either as total market capitalization or as the book value of total assets. Crawford et al. (2012) explain that firm size proxies for "various dimensions of the firm's information environment, including media exposure and the overall level of investor interest." (p. 1536). At the same time, firm size is also related to complexity, as large firms tend to have more complex organizational structures (Bushman, Piotroski, and Smith, 2004), more intangibles on their balance sheets (Gu and Wang, 2005), and face stricter internal control reporting requirements than small firms (Ettredge, Johnstone, Stone, and Wang, 2011), resulting in longer financial reports. Other variables positively related to size, such as earnings (i.e.; return on equity), market share, institutional ownership, and analyst coverage are also used frequently to proxy for information production.

We calculate firm size *(QSIZE$_{it}$)* as the natural logarithm of the product of quarter-end closing share prices times the number of common and ordinary shares outstanding and take its first difference ($\Delta QSIZE_{it}$). Although, our sample has a slight tilt towards large firms, it still includes substantial size variation. The average *QSIZE$_{it}$* is 5.942 (corresponding to $381 million), while the variable ranges between $8.6 million and $52.1 billion. Book to market (*QBTM$_{it}$*) is calculated as the natural logarithm of the book value of equity divided by the market value of

equity. Leverage (*LEVERAGE$_{it}$*) is calculated as total liabilities divided by total assets. Average leverage of our sample is 0.438.

Return on Equity *(ROE$_{it}$)* is the firm's earnings scaled by the book value of equity in a given quarter. Average quarterly *ROE$_{it}$* is 1.4%. Quarterly changes in ROE are denoted as *ΔROE$_{it}$*. We include this variable to control for the effect of performance on comovement, as in Hutton et al. (2009).

Similar to firm characteristics, analyst coverage is often used to proxy for information production as analysts' reports and forecasts provide information (Lys and Soo, 1995; Lang, Lins, and Miller, 2003). At the same time, several studies find that analysts gravitate towards large firms (Brennan and Hughes, 1991), firms that disclose more (Bhushan, 1989; Lang and Lundholm, 1996), and firms with less readable financial reports (Lehavy et al., 2011). We measure the intensity of analyst coverage *(ANALYSTS$_{it}$)* following Piotroski and Roulstone (2004) as the logarithmic transformation (*ln*[1+*x*]) of the number of earnings estimates on the stock in a given quarter. The mean and median number of analyst forecasts in our sample is four, while the range is between zero and 26.

An additional caveat with the use of traditional information production variables in our regressions is that they might be positively correlated with 10-Q word counts and with one another. For this reason, even though we include them in our analysis for the sake of completeness, we refrain from conducting inference on their coefficient estimates. Instead, we introduce two variables uncorrelated to changes in complexity: an indicator that takes the value of one when a news-grabbing event coincides with the complexity shock (*INFO$_{it}$)*, and changes in equity ownership of dedicated financial institutions (*DED_IOPERC$_{it}$)*.

INFO$_{it}$ is our proxy for news grabbing events and is an indicator variable that takes the value of one if in the three months leading up to the quarterly filing there was an equity issuance, a debt issuance, a stock split, or a special dividend. If none of these events took place during the three months before the filing, *INFO$_{it}$* takes a value of zero. Stock splits and special dividends are obtained directly from the CRSP events file. Equity issuances are detected through changes in the number of shares issued. Debt issuances are approximated by the 10% largest long term debt increases experienced by firms during the sample period.

Dedicated institutional ownership (*DED_IOPERC$_{it}$*) is measured as the percentage of the firm's shares owned by dedicated financial institutions. According to Bushee and Noe (2000) and Bushee (2001), dedicated financial institutions are those that invest in long-term, stable firms and seek longer-term dividend income and capital appreciation. These institutions are characterized by large portfolio allocations, low turnovers, and relationship investing. Because of their high levels of specialization, dedicated institutions are expected to trade on private information more often than other financial institutions, particularly index funds. Therefore, we expect that price fluctuations from stocks owned more heavily by dedicated institutions will reveal more firm-specific information than those from stocks owned by other types of institutional investors. We obtain quarterly institutional ownership data from the Thomson Financial 13(f) database as of the end of each fiscal quarter. Δ*DED_IOPERC$_{it}$* is the change in dedicated institutional ownership between the quarter before and the quarter of the filing. Average dedicated institutional ownership in our sample is 4.2% (median of 0%) and ranges between 0% and 33.9%.

Another control variable to account for the effect of the information environment on peer correlations is opacity. As explained in the introduction, opacity and complexity are separate

attributes of information that should produce higher comovement. The theoretical model in Jin and Myers (2006) explains that opacity provides corporate insiders with opportunities to extract private control benefits and suppress the release of firm-specific information. The lack of firm-specific information causes stocks to correlate. We follow Hutton et al. (2009) and use discretionary accruals to capture opacity ($OPACITY_{it}$), where discretionary accruals are defined as the residuals from a performance controlled Jones (1991) regression as outlined in Kothari, Leone, and Wasley (2005). As expected, average (median) $OPACITY_{it}$ in the sample is near zero, at 0.2% (0.1%) of total assets, and there is significant variation with a range between -12.5% and 13.8% of total assets. $\Delta OPACITY_{it}$ is the quarterly change in $OPACITY_{it}$

In addition to capturing increases in media attention as firms increase their sales revenue and net income at the expense of other firms in their industry, thus reducing comovement, changes in market share and earnings may also account for intra-industry information transfers between firms experiencing positive performance and other firms in the same industry (Foster, 1981). As a firm becomes dominant in an industry, more investors will use its news as industry information, thus increasing comovement (Hou, 2007). We calculate market share (MKT_SHARE_{it}) as the percentage of firm sales on total industry sales in a given quarter, where industry is defined as all firms with the same four-digit SIC code. While most of the sample firms control a small portion of their markets, several of them could be dominant enough to be considered bellwether firms. The average (median) market share is 7.4% (1.2%), while the variable ranges between near zero and 89.6%. Quarterly market share change is denoted as ΔMKT_SHARE_{it}.

For completeness, our regressions also include institutional ownership of non-dedicated financial institutions ($INSTPERC_{it}$). These financial institutions are less likely to trade on firm-

specific information because they follow either indices or speculative strategies. Analogous to the calculation for DED_IOPERC_{it}, we measure $INSTPERC_{it}$ as the proportion of shares held by non-dedicated financial institutions to total shares outstanding, again following Bushee (2001). Mean and median $INSTPERC_{it}$ is 33% and 26.5%, respectively. Our sample contains stocks with a full range of institutional ownerships, from zero to 100%.

3.5.2 Transaction Costs

In general, an increase in transaction costs increases investors' underreaction to firm-specific news and reduces comovement. This happens because transaction costs delay the occurrence of profitable trades by arbitrageurs that trade on new information. Amihud and Mendelson (1987) and Damodaran (1993) posit that stock prices adjust to news at different speeds because not all investors receive information signals at the same time, so the trades from late-informed investors lag behind the trades from early-informed investors (e.g., Froot, Scharfstein, and Stein, 1993; Hirshleifer, Subrahmanyam, and Titman, 1994). Market frictions exacerbate differences in information arrival and trade executions between early and late-informed investors, delaying price adjustment. Hence, correlations across stock returns should increase as cross-sectional differences in information arrival and trade execution subside. We expect a negative relation between peer comovement and market frictions. We utilize two market frictions proxies, the first-order autocorrelation coefficient of stocks returns and share turnover.

According to Black (1986) the noise component of price fluctuations *(NOISE$_{it}$)* has a temporary effect on prices. Given the temporary nature of noise, we expect it to contribute to the negative autocorrelation of returns (Amihud and Mendelson, 1987). We measure noise *(NOISE$_{it}$)* following Amihud and Mendelson (1987) as the first-order autocorrelation coefficient of daily

returns of stock i during quarter t. We multiply $NOISE_{it}$ times minus one for ease of interpretation; this way, we obtain a measure that increases when noise increases and declines when noise declines. We compute the change in $NOISE_{it}$ between the three month period before the 10-Q filing month and the three month period after the 10-Q filing month. To remain consistent with the measurement of return synchronicity ($\Delta SYNCH_{it}$), we exclude the 10-Q filing month from both estimation periods. We denote the change in noise around the 10-Q filing month as $\Delta NOISE_{it}$.

In addition to including noise as an aggregate market frictions measure, we include stock turnover *(TURNOVER_{it})*, a measure of transaction costs. $TURNOVER_{it}$ is measured as dollar volume divided by number of shares outstanding at the end of the quarter. We calculate turnover change *($\Delta TURNOVER_{it}$)* between the three month period before the 10-Q filing month and the three month period after the 10-Q filing month.

3.5.3   Other correlated effects

Joint changes in firm fundamentals (3) and joint underreaction to firm-specific news (4) are controlled for by using industry indicator variables. We also note that the use of industry indicators would eliminate the effect of joint increases in complexity caused by changes in industry-specific accounting rules and disclosure standards. This means that the coefficients of our treatment variables might understate the total effect of complexity shocks on complementarities. Similarly, trends in information production, transaction costs and firm fundamentals (5) are controlled for using time indicator variables. Just like with the industry indicators the use of time indicators will also eliminate the effect of trends in financial reporting complexity not captured by 10-Q word counts.

4. Hypothesis testing

Table 5 summarizes the Pearson correlation coefficients for the variables used in our analysis.

Panel A of Table 5 presents correlation coefficients of the levels of the variables in our analysis.

The sign and magnitude of these coefficients are generally consistent with prior literature. Panel

B of Table 5 presents correlation coefficients for these variables' first-differences, which is the

way these variables are introduced in our regressions. Consistent with our hypothesis, Panel B of

Table 5 shows a positive and significant correlation between change in complexity

($\Delta COMPLEX_{it}$) and change in peer comovement ($\Delta SYNCH_{it}$) of 0.02. This value corresponds to a

correlation in the levels of the variables of 0.28. Correlation coefficients for levels and first-

differences are statistically significant at less than the 1% level. Also consistent with our

hypothesis, change in average peer complexity ($\Delta AVG\_PR\_COMPLEX_{it}$) is positive (0.01) and

statistically significant (p-value=0.03). While this is not a direct test of our hypothesis, this

coefficient is also consistent with the theory that increases in complexity are associated with

increases in comovement with peers. A complexity shock either on the firm of interest or on its

peers should stimulate the acquisition of low cost signals, resulting in higher comovement.

Table 5
Pooled correlation for variables used to determine the influence of complexity on peer stock comovement

This table reports pairwise Pearson correlations for the variables used in our analysis. Panel A shows correlations of the variables' levels and Panel B shows correlations of the variables' first-differences. $\Delta SYNCH_{it}$ is the change in peer group synchronicity around the filing date of the 10-Q, where synchronicity is measured as a log transformation of the R^2 from a regression of firm returns on value-weighted peer group returns. $\Delta COMPLEX_{it}$ is the change from last quarter of the natural logarithm of the number of words in the 10-Q report. All other variables are described in detail in the Appendix.

Panel A: Correlations of levels

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	$SYNCH_{it}$	1.00														
2	$COMPLEX_{it}$	0.28	1.00													
		(0.00)														
3	$AVG\_PR\_COMPLEX_{it}$	0.29	0.47	1.00												
		(0.00)	(0.00)													
4	$QSIZE_{it}$	0.44	0.37	0.2	1.00											
		(0.00)	(0.00)	(0.00)												
5	MKT_SHARE_{it}	0.1	0.05	-0.02	0.37	1.00										
		(0.00)	(0.00)	(0.00)	(0.00)											
6	$ANALYSTS_{it}$	0.4	0.37	0.21	0.76	0.21	1.00									
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)										
7	$INSTPERC_{it}$	0.35	0.31	0.31	0.46	0.13	0.47	1.00								
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)									
8	$QBTM_{it}$	-0.09	-0.05	0	-0.36	-0.02	-0.24	-0.07	1.00							
		(0.00)	(0.00)	(0.85)	(0.00)	(0.00)	(0.00)	(0.00)								
9	ROE_{it}	0.08	-0.08	-0.04	0.22	0.14	0.11	0.1	-0.05	1.00						
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)							
10	$LEVERAGE_{it}$	0.08	0.15	0.04	0.22	0.23	0.11	0.08	-0.04	0.12	1.00					
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)						
11	DED_IOPERC_{it}	0	-0.08	-0.16	0.11	0.04	0.11	0.15	-0.02	0.04	0.04	1.00				
		(0.42)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)					
12	$INFO_{it}$	-0.03	0	-0.07	0	-0.02	-0.02	-0.05	-0.11	-0.06	0.01	-0.01	1.00			
		(0.00)	(0.92)	(0.00)	(0.49)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.02)	(0.18)				
13	$OPACITY_{it}$	-0.04	-0.02	-0.03	-0.07	-0.01	-0.06	-0.05	0.05	-0.06	0	-0.02	0.06	1.00		
		(0.00)	(0.00)	(0.00)	(0.00)	(0.05)	(0.00)	(0.00)	(0.00)	(0.00)	(0.75)	(0.00)	(0.00)			
14	$NOISE_{it}$	-0.1	-0.13	-0.05	-0.22	-0.05	-0.2	-0.11	0.15	0	-0.04	-0.02	-0.05	-0.01	1.00	
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.73)	(0.00)	(0.01)	(0.00)	(0.37)		
15	$TURNOVER_{it}$	0.25	0.26	0.17	0.25	-0.02	0.36	0.28	-0.2	-0.04	-0.07	-0.04	0.09	-0.01	-0.19	1.00
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.36)	(0.00)	

Panel B: Correlations of changes

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	ΔSYNCH$_{it}$	1.00														
2	ΔCOMPLEX$_{it}$	0.02	1.00													
		(0.00)														
3	ΔAVG_PR_COMPLEX$_{it}$	0.01	0.14	1.00												
		(0.03)	(0.00)													
4	ΔQSIZE$_{it}$	0.06	0.01	0.01	1.00											
		(0.00)	(0.12)	(0.02)												
5	ΔMKT_SHARE$_{it}$	0.01	-0.03	-0.03	0.01	1.00										
		(0.36)	(0.00)	(0.00)	(0.02)											
6	ΔANALYSTS$_{it}$	0.02	0.05	0.08	0.08	0.00	1.00									
		(0.00)	(0.00)	(0.00)	(0.00)	(0.62)										
7	ΔINSTPERC$_{it}$	0.03	-0.01	-0.01	0.20	0.01	0.05	1.00								
		(0.00)	(0.12)	(0.18)	(0.00)	(0.29)	(0.00)									
8	ΔQBTM$_{it}$	-0.05	-0.01	-0.02	-0.84	0.01	-0.02	-0.14	1.00							
		(0.00)	(0.20)	(0.00)	(0.00)	(0.03)	(0.01)	(0.00)								
9	ΔROE$_{it}$	0.00	0.00	0.02	0.08	0.09	0.01	0.02	0.04	1.00						
		(0.82)	(0.53)	(0.00)	(0.00)	(0.00)	(0.15)	(0.01)	(0.00)							
10	ΔLEVERAGE$_{it}$	-0.01	0.01	-0.01	-0.12	0.04	-0.05	-0.05	-0.18	-0.15	1.00					
		(0.20)	(0.08)	(0.12)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)						
11	ΔDED_IOPERC$_{it}$	-0.03	0.01	0.04	-0.01	-0.01	0.01	-0.34	0.01	0.01	-0.02	1.00				
		(0.00)	(0.16)	(0.00)	(0.03)	(0.22)	(0.04)	(0.00)	(0.03)	(0.12)	(0.01)					
12	INFO$_{it}$	0.00	0.00	-0.04	0.15	0.04	0.07	0.07	0.01	-0.01	0.05	-0.03	1.00			
		(0.45)	(0.49)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)	(0.24)	(0.00)	(0.00)				
13	ΔOPACITY$_{it}$	0.00	0.01	0.01	0.03	0.00	0.01	0.01	0.00	-0.01	0.00	0.00	0.07	1.00		
		(0.96)	(0.37)	(0.22)	(0.00)	(0.94)	(0.27)	(0.03)	(0.43)	(0.12)	(0.83)	(0.80)	(0.00)			
14	ΔNOISE$_{it}$	-0.01	0.00	-0.01	0.01	0.00	0.01	0.00	-0.01	0.01	0.00	0.01	0.00	0.00	1.00	
		(0.21)	(0.73)	(0.20)	(0.13)	(0.63)	(0.12)	(0.58)	(0.22)	(0.28)	(0.74)	(0.05)	(0.42)	(0.63)		
15	ΔTURNOVER$_{it}$	0.03	-0.01	-0.03	0.15	0.01	0.02	0.08	-0.13	-0.01	0.00	-0.02	0.04	0.00	0.03	1.00
		(0.00)	(0.17)	(0.00)	(0.00)	(0.37)	(0.01)	(0.00)	(0.00)	(0.17)	(0.99)	(0.01)	(0.00)	(0.47)	(0.00)	

We test our hypothesis in a multivariable setting by regressing changes in firm complexity on changes in peer comovement while controlling for changes in the complexity of the firm's peers, information production, market frictions, and fixed effects. We fit the following pooled regression:

$$\Delta SYNCH_{it} = \beta_0 + \beta_1 \Delta COMPLEX_{it} + \beta_2 \Delta AVG\_PR\_COMPLEX_{it} + \beta_3 \Delta QSIZE_{it} + \beta_4 \Delta MKT\_SHARE_{it} + \beta_5 \Delta ANALYSTS_{it} + \beta_6 \Delta INSTPERC_{it} + \beta_7 \Delta QBTM_{it} + \beta_8 \Delta ROE_{it} + \beta_9 \Delta LEVERAGE_{it} + \beta_{10} \Delta DED\_IOPERC_{it} + \beta_{11} \Delta INFO_{it} + \beta_{12} \Delta OPACITY_{it} + \beta_{13} \Delta NOISE_{it} + \beta_{14} \Delta TURNOVER_{it} + \varepsilon_{it},$$ (3)

We expect a positive coefficient on $\Delta COMPLEX_{it}$ (i.e., $\beta_1 > 0$) from equation (3). Control variables were discussed in section 3, and are formally defined in the Appendix. We include industry and year indicators to capture any unobservable industry and year fixed-effects. Industry indicator variables correspond to two-digit SIC codes. The F-test statistics for the joint significance of industry and year effects are 25.24 and 1.77, respectively, both statistically significant at the 1% level, thus justifying the use of two-way fixed effects. We use heteroskedasticity-robust standard errors, clustered by firm.

Table 6 reports regression coefficients and their respective two-tailed *p*-values. Model (1) in Table 6 presents the baseline regression coefficient for $\Delta COMPLEX_{it}$ in a pooled regression without controls or fixed-effects. Model (2) in Table 6 presents regression coefficients for the pooled ordinary least squares regression without fixed-effects. Model (3) in Table 6 presents regression coefficients for the pooled ordinary least squares regression with year and industry indicators. Model (4) in Table 6 presents regression coefficients for the pooled ordinary least squares regression with year-industry indicators. This last specification is included to account for the potential presence of time-industry effects resulting from the implementation of accounting standards that could produce industry-specific increases in 10-Q lengths in certain periods and return synchronicity.

Table 6
Influence of complexity on peer stock comovement - warranted multiples selection

$$\Delta SYNCH_{it} = \beta_0 + \beta_1\Delta COMPLEX_{it} + \beta_2\Delta AVG\_PR\_COMPLEX_{it} + \beta_3\Delta QSIZE_{it} + \beta_4\Delta MKT\_SHARE_{it} + \beta_5\Delta ANALYSTS_{it} + \beta_6\Delta INSTPERC_{it}$$
$$+\beta_7\Delta QBTM_{it} + \beta_8\Delta ROE_{it} + \beta_9 LEVERAGE_{it} + \beta_{10}\Delta DED\_IOPERC_{it} + \beta_{11}\Delta INFO_{it} + \beta_{12}\Delta OPACITY_{it} + \beta_{13}\Delta NOISE_{it} + \beta_{14}\Delta TURNOVER_{it} + \varepsilon_{it},$$

This table presents coefficient estimates from regressions of $\Delta SYNCH_{it}$ on $\Delta COMPLEX_{it}$. Peer groups are identified following the method in Bhojraj and Lee (2002). $\Delta SYNCH_{it}$ is the change in peer group synchronicity around the filing date of the 10-Q, where synchronicity is measured as a log transformation of the R^2 from a regression of firm returns on value-weighted peer group returns. $\Delta COMPLEX_{it}$ is the change from last quarter of the natural logarithm of the number of words in the 10-Q report. Two tailed p-values (to the right of coefficient estimates) are calculated using heteroskedasticity-robust standard errors clustered by firm. Industry effects are included using two-digit SIC codes. See the Appendix for a description of control variables.

Variable	(1) $\Delta SYNCH_{it}$		(2) $\Delta SYNCH_{it}$		(3) $\Delta SYNCH_{it}$		(4) $\Delta SYNCH_{it}$	
$\Delta COMPLEX_{it}$	0.1016	(0.001)	0.0936	(0.003)	0.0964	(0.002)	0.0995	(0.002)
$\Delta AVG\_PR\_COMPLEX_{it}$			0.0762	(0.171)	0.0836	(0.139)	0.0954	(0.100)
$\Delta QSIZE_{it}$			0.4820	(0.000)	0.4390	(0.000)	0.4357	(0.000)
ΔMKT_SHARE_{it}			0.8643	(0.263)	0.6630	(0.390)	0.8237	(0.292)
$\Delta ANALYSTS_{it}$			0.1333	(0.048)	0.1504	(0.026)	0.1504	(0.029)
$\Delta INSTPERC_{it}$			0.5264	(0.070)	0.5934	(0.042)	0.5683	(0.057)
$\Delta QBTM_{it}$			0.0065	(0.952)	-0.0624	(0.567)	-0.0575	(0.604)
ΔROE_{it}			-0.1530	(0.538)	-0.1429	(0.567)	-0.1858	(0.462)
$\Delta LEVERAGE_{it}$			0.0032	(0.993)	-0.1013	(0.792)	-0.1340	(0.731)
ΔDED_IOPERC_{it}			-1.8085	(0.001)	-2.0319	(0.000)	-2.0552	(0.001)
$INFO_{it}$			-0.1136	(0.014)	-0.0986	(0.034)	-0.1070	(0.023)
$\Delta OPACITY_{it}$			-0.1001	(0.778)	0.0769	(0.833)	0.1036	(0.778)
$\Delta NOISE_{it}$			-0.1072	(0.201)	-0.1338	(0.114)	-0.1223	(0.155)
$\Delta TURNOVER_{it}$			0.0364	(0.009)	0.0272	(0.052)	0.0254	(0.075)
CONSTANT	0.0239	(0.081)	0.0227	(0.249)	0.0977	(0.843)	-0.4240	(0.000)
YEAR FIXED EFFECTS	NO		NO		YES	(14)	NO	
INDUSTRY FIXED EFFECTS	NO		NO		YES	(54)	NO	
INDUSTRY/YEAR FIXED EFFECTS	NO		NO		NO		YES	(581)
N	26,484		26,484		26,484		26,484	
Adj R-squared	0.001		0.006		0.012		0.008	

Consistent with our hypothesis, there is a positive and significant coefficient for $\Delta COMPLEX_{it}$ in all model specifications. This coefficient remains remarkably unchanged after the introduction of control variables, year, industry, and year-industry effects. According to Model (3), the coefficient on $\Delta COMPLEX_{it}$ is 0.0964 (*p*-value0.002). This implies that holding all other variables constant, a one-standard-deviation increase in complexity induces an increase in R^2 (at the mean) of 4.97%. This is a sizable amount if one considers that the average R^2 for the sample is 10.60%, and the median is 4.50%. This coefficient suggests that complexity has a significant effect on stock correlations.[21]

The coefficient estimate for change in average peer firm complexity ($\Delta AVG\_PR\_COMPLEX_{it}$) is positive, and only statistically significant from zero in Model (4), at the 10 percent level (however, using one-tailed tests the coefficients are statistically significant across all specifications). The coefficient estimate for increases in firm size ($\Delta QSIZE_{it}$) is positive and significant (0.4390, *p*-value<0.001). As discussed above, the sign on this variable is difficult to interpret because size proxies for many firm attributes (including complexity and the information environment which have competing predictions). That said, the results indicate that as firms grow larger, they tend to comove more strongly with their peers.

Coefficient estimates for the change in market share (ΔMKT_SHARE_{it}), change in return on equity *(ΔROE_{it})*, change in the book-to-market ratio ($\Delta QBTM_{it}$), change in leverage *($\Delta LEVERAGE_{it}$)*, change in opacity *($\Delta OPACITY_{it}$)*, and change in noise *($\Delta NOISE_{it}$)* are statistically insignificant.

---

[21] Our predictions center on unexpected increases in reporting complexity. For completeness we include the entire distribution of $\Delta COMPLEX_{it}$ in our analysis. However, it is not necessarily clear that decreases in complexity should translate into decreases in comovement. In untabulated analysis we find stronger economic significance when we exclude decreases in $\Delta COMPLEX_{it}$ from our analysis. Thus, the effects documented may understate the effect of unexpected complexity on peer return comovement.

The coefficient estimate for $\Delta ANALYSTS_{it}$ is positive (0.1504, p-value=0.026). This implies that increases in analyst coverage on a stock results in increases in the correlation between a stock and its peers. Again, we caution the reader on the interpretation of these variables given that analysts tend to be attracted to more complex firms (Lehavy et al., 2011) and size, market share, performance, book-to-market, leverage, and analyst following are also associated with information production.

Turning to our information production proxies we find that regression coefficients for dedicated institutional ownership (ΔDED_IOPERC_{it}) and the presence of news grabbing events at the time of the complexity shock ($INFO_{it}$) are both negative and statistically significant at conventional levels. The coefficient estimate for dedicated institutional ownership is -2.0319 (p-value<0.001). This indicates that a one-standard-deviation increase in the proportion of shares owned by dedicated financial institutions results in a reduction in peer comovement of 5.32%. These results are in line with Veldkamp's (2006) prediction that increases in information production disincentivize low-cost information acquisition and reduce comovement. The coefficient estimate for non-dedicated institutional ownership ($\Delta INSTPERC_{it}$), is positive and significant (0.5934 p-value=0.042). The opposite signs of the regression coefficients for dedicated financial institutions (negative) and non-dedicated financial institutions (positive) are also consistent with information choice theory, and with Bushee's (2001) institutional investor classification. As dedicated institutional ownership is more likely to increase demand for firm-specific information than other investors (i.e., indexers and speculators), the presence of dedicated institutions should reduce comovement, while the presence of non-dedicated institutions should increase it.

Finally, we observe that increases in firm turnover, the result of reductions in market frictions, lead to increases in peer correlations. This could indicate a more synchronized response of prices to new information as market frictions subside and prices become more informative.

5. Additional Analysis

5.1 Random Peer Assignment

As explained in Section 3, we construct our sample following Bhojraj and Lee (2002), in an attempt to mimic the peer selection process of an analyst or investor that identifies peers by comparing financial ratios and reviewing soft information. However, there is a concern that the extensive data requirements in Bhojraj and Lee (2002) might have biased our sample towards large stocks with naturally high comovement with other stocks. If that were the case our Table 6 results could be the result of sample selection bias and not be representative of the economic relation between complexity and comovement. To make sure that our results are robust to alternative peer matching procedures, we devise an alternative, and much simpler, peer group formation method by choosing at random four stocks from the firm's two-digit SIC industry code to serve as peers. These firms will share the firm of interest's industry affiliation but could have other characteristics when it comes to size, age, or level of horizontal or vertical integration. This new matching eliminates many of the data restrictions that limited the size of our sample. The new sample consists of 43,231 firm-quarter observations, almost twice as many observations as the original sample. We fit the regression in Equation (3) on the new sample and find results consistent with a positive relationship between increases in complexity and increases in comovement.

37

Consistent with Table 6, the coefficients on $\varDelta COMPLEX_{it}$ in Table 7 are positive. However, statistical significance at conventional levels is only achieved in one-tailed tests, except in Model (1) where no control variables are included (two-tailed tests tabulated). Also consistent with Table 6, the coefficient on changes in average peer firm complexity ($\varDelta AVG\_PR\_COMPLEX_{it}$) is positive (0.3650), but in this specification is highly statistically significant (p-value<0.001). These results suggest that when using random peer assignment the dominant effect appears to be through increases in average peer firm complexity, perhaps due to the less precise peer assignment process. Overall, it is not just that complexity shocks in the firm of interest result in higher peer comovement, it is also complexity shocks affecting the firm's peers as well. A one-standard-deviation increase in the change in average peer complexity induces a change in R^2 (at the mean) of 11.66%. Taken together, these results provide additional evidence consistent with our hypothesis that investors respond to complexity shocks by acquiring low-cost information, which generates excess comovement.

Table 7
Influence of complexity on peer stock comovement - random within-industry selection

$$\Delta SYNCH_{it} = \beta_0 + \beta_1 \Delta COMPLEX_{it} + \beta_2 \Delta AVG\_PR\_COMPLEX_{it} + \beta_3 \Delta QSIZE_{it} + \beta_4 \Delta MKT\_SHARE_{it} + \beta_5 \Delta ANALYSTS_{it} + \beta_6 \Delta INSTPERC_{it}$$
$$+\beta_7 \Delta QBTM_{it} + \beta_8 \Delta ROE_{it} + \beta_9 LEVERAGE_{it} + \beta_{10} \Delta DED\_IOPERC_{it} + \beta_{11} \Delta INFO_{it} + \beta_{12} \Delta OPACITY_{it} + \beta_{13} \Delta NOISE_{it} + \beta_{14} \Delta TURNOVER_{it} + \varepsilon_{it}$$

This table presents coefficient estimates from regressions of $\Delta SYNCH_{it}$ on $\Delta COMPLEX_{it}$. Peer groups are four firms chosen at random within the firm's two-digit SIC industry. $\Delta SYNCH_{it}$ is the change in peer group synchronicity around the filing date of the 10-Q, where synchronicity is measured as a log transformation of the R^2 from a regression of firm returns on value-weighted peer group returns. $\Delta COMPLEX_{it}$ is the change from last quarter of the natural logarithm of the number of words in the 10-Q report. Two tailed p-values (to the right of coefficient estimates) are calculated using heteroskedasticity-robust standard errors clustered by firm. Industry effects are included using two-digit SIC codes. See the Appendix for a description of control variables.

Variable	(1) $\Delta SYNCH_{it}$		(2) $\Delta SYNCH_{it}$		(3) $\Delta SYNCH_{it}$		(4) $\Delta SYNCH_{it}$	
$\Delta COMPLEX_{it}$	0.0732	(0.004)	0.0382	(0.136)	0.0417	(0.106)	0.0413	(0.114)
$\Delta AVG\_PR\_COMPLEX_{it}$			0.3286	(0.000)	0.3650	(0.000)	0.5307	(0.000)
$\Delta QSIZE_{it}$			0.4629	(0.000)	0.4192	(0.000)	0.3925	(0.000)
ΔMKT_SHARE_{it}			-0.7955	(0.317)	-0.9313	(0.242)	-0.7421	(0.357)
$\Delta ANALYSTS_{it}$			-0.1446	(0.008)	-0.1143	(0.038)	-0.1141	(0.039)
$\Delta INSTPERC_{it}$			0.6738	(0.008)	0.5861	(0.023)	0.5989	(0.020)
$\Delta QBTM_{it}$			0.1055	(0.163)	0.0555	(0.464)	0.0555	(0.470)
ΔROE_{it}			0.0829	(0.415)	0.0821	(0.423)	0.0712	(0.489)
$\Delta LEVERAGE_{it}$			0.4362	(0.109)	0.2864	(0.292)	0.2275	(0.404)
ΔDED_IOPERC_{it}			-2.4605	(0.000)	-2.6108	(0.000)	-2.7690	(0.000)
$INFO_{it}$			-0.1114	(0.002)	-0.1004	(0.006)	-0.1041	(0.005)
$\Delta OPACITY_{it}$			0.4676	(0.080)	0.3850	(0.168)	0.4155	(0.138)
$\Delta NOISE_{it}$			-0.1352	(0.038)	-0.1707	(0.009)	-0.1511	(0.022)
$\Delta TURNOVER_{it}$			0.0522	(0.000)	0.0470	(0.000)	0.0506	(0.000)
CONSTANT	0.0519	(0.000)	0.0353	(0.039)	0.2519	(0.621)	3.1446	(0.000)
YEAR FIXED EFFECTS	NO		NO		YES	(14)	NO	
INDUSTRY FIXED EFFECTS	NO		NO		YES	(62)	NO	
INDUSTRY/YEAR FIXED EFFECTS	NO		NO		NO		YES	(636)
N	43,231		43,231		43,231		43,231	
Adj R-squared	0.0002		0.0062		0.0121		0.0211	

5.2 Changes in 10-K Length

Because we are interested in unexpected increases in financial reporting complexity, our primary analysis focuses on the shortest window of SEC reporting (quarterly reports). Alternatively, annual reports could be used. However, since a great deal of reporting activity occurs between each 10-K filing, we expect that annual reports will be a weaker setting to test our prediction. In addition, we also expect the information environment surrounding 10-K filings to be stronger, relative to 10-Q filings. Consistent with these expectations, substituting changes in 10-K length as our measure of financial reporting complexity and examining changes in comovement of the peer portfolio around the 10-K filing produces insignificant results. Overall, this suggests that changes in 10-K length either are not unexpected, and/or are more easily remedied by a more robust information environment and therefore do not produce the same consequences as changes in 10-Q length.

6. Conclusion

We use information choice principles to explain investors' reaction to complexity shocks (e.g., Veldkamp, 2006; Chamley, 2007; Ganguli and Yang, 2009; Manzano and Vives, 2011; Mele and Sangiorgi, 2015). We propose and show that investors respond to complexity shocks in quarterly reports by increasing demand for low-price information on the stock. Consistent with our hypothesis, we find a positive relation between changes in 10-Q word counts, our complexity proxy, and changes in 1) the number of internet searches on the stock and 2) the R^2 of the regression of stock returns on the average returns of its four closest peers. These results arise because, consistent with Veldkamp's (2006) prediction, return comovement increases when investors acquire more low-price information signals. We observe this positive relation for

complexity shocks in the firm of interest and find some evidence that this positive relation also exists for complexity shocks in its strategic peers.

Throughout our tests, we take great care to differentiate between comovement increases caused by complexity shocks and comovement increases caused by lower information production and higher market frictions. Isolating the effect of complexity shocks on comovement from the effect of changes in information production pose a significant challenge because most of the traditional information production measures in the literature are positively related to complexity. For this reason, we introduce two alternative information production proxies: institutional ownership of dedicated financial institutions and an indicator that takes the value of one when the complexity shock coincides with a corporate news-grabbing event. The regression coefficients from these variables are consistent with prior literature and show a negative relation between comovement and information production. This suggests that the effect of complexity shocks on the firm's information environment can be, at least partially, offset by the production of news and analysis from other sources. In this regard, the role of dedicated financial institutions to reveal firm-specific information through their trades seems particularly important.

One limitation of our study, that is a limitation of all studies using a length based proxy for financial reporting complexity, is that the exact source of underlying complexity that investors face cannot be precisely identified. As Loughran and McDonald note, complexity and readability cannot really be disentangled (2014, pg. 1646). Future research in this area could work to disentangle these effects.

Our study contributes by providing empirical evidence that complexity generates excess comovement. This seems to indicate that investors see complexity shocks and reductions in transparency as equivalent. So far, existing studies had only demonstrated that comovement

increased in response to reductions in the number of information signals (Jin and Myers, 2006; Hutton et al., 2009; Brockman, Liebenberg, and Schutte, 2010; Hameed, Morck, Shen, and Yeung, 2015).

Although we take a modest first step at incorporating information choice concepts into the broad accounting narratives literature, the results from our study might have broader implications for other actively researched areas of Accounting and Finance. For instance, even though we obtain evidence suggesting that investors respond to more complex accounting reports by demanding information from other sources, we do not know with certainty how effective this strategy is at mitigating uncertainty. Predictions from general equilibrium models with endogenous information acquisition (Clinch and Lombardi, 2011; Opp, 2015; Han and Yang, 2013) are contradictory and highly dependent on model assumptions. This leaves managers and regulators with no definite answer on how complexity influences cost of capital. The use of measures derived from quantitative and qualitative aspects of accounting narratives could prove beneficial as we test these models empirically.

Finally, our results might also prove useful to understand the economic importance of accounting information complexity on market efficiency. Users, preparers and regulators have highlighted the importance of accounting information's attributes for a well-functioning capital market, with obvious consequences for the economy at large. However, if investors are able to stimulate the release of private information from other sources by increasing their demand, it is possible that the effects of complexity on market efficiency might be smaller than anticipated. Additional studies that link measures derived from accounting narratives with market efficiency proxies could help address this question. Some examples of possible areas of future research could include studying how changes in the reporting environment impacts comovement among

firms, such as Regulation Fair Disclosure, the Plain English Initiative, and the Sarbanes Oxley Act.

Acknowledgements

Appendix. Variable definitions

This appendix reports details about the calculation of variables used in our analysis. All continuous variables are winsorized at 1% and 99% to control for outliers. Unless otherwise noted, variable names identified in the definition column correspond to the quarterly COMPUSTAT Xpressfeed data set or the CRSP data set.

Variable	Definition
Peer Stock Comovement Variables	
R^2	The coefficient of determination from estimation of the following model: $RET_{it} = \beta_1 + \beta_2 PEER\_RET_{it} + \varepsilon_{it}$ where RET is the returns for firm i, and PEER_RET are value-weighted returns for the 4 closest peer firms identified during the fiscal year. The model is estimated using daily returns, for each of the two following periods: (1) the three month period before the filing of the 10-Q and (2) the three month period following the filing of the 10-Q. The month of filing of the 10-Q is excluded from both periods.
$SYNCH_{it}$	$\ln(R^2/(1-R^2))$ for the three month period following the filing of the 10-Q.
$\Delta SYNCH_{it}$	The change in peer group synchronicity around the filing of the 10-Q.
Financial Reporting Complexity Variables	
RAW_WORDS_{it}	The number of words in the 10-Q.
$COMPLEX_{it}$	The natural logarithm of the number of words in the 10-Q.
$\Delta COMPLEX_{it}$	The change in COMPLEX from the prior 10-Q.
$AVG\_PR\_COMPLEX_{it}$	The mean of COMPLEX for the members of the identified peer group.
$\Delta AVG\_PR\_COMPLEX_{it}$	The change in AVG_PR_COMPLEX.
Other Control Variables for Tests of Influence of Complexity on Peer Stock Comovement	
$ANALYSTS_{it}$	The natural logarithm of one plus the number of individual analyst earnings estimates as of the end of the quarter.
$\Delta ANALYSTS_{it}$	The quarterly change in ANALYSTS.
DED_IOPERC_{it}	The percentage of dedicated institutional ownership as classified in Bushee (2001).
ΔDED_IOPERC_{it}	The quarterly change in DED_IOPERC.
$INFO_{it}$	An indicator for the presence of any of the following events: (1) Debt issuance proxied by a firm's change in long-term debt being in the top decile, (2) Equity issuance proxied by a firm's change in shares issued being in the top decile, (3) a stock split in the 3 month's leading up to the filing of the 10-Q, or (4) a special dividend issuance in the 3 month's leading up to the filing of the 10-Q, zero otherwise.
$INSTPERC_{it}$	The percentage of institutional ownership measured as the sum of non-dedicated shares/shares outstanding. Where non-dedicated shares are all institutional shares held at the end of the quarter, except dedicated shares as classified in Bushee (2001).
$\Delta INSTPERC_{it}$	The quarterly change in INSTPERC.
$LEVERAGE_{it}$	Total liabilities scaled by total assets (ltq/atq).
$\Delta LEVERAGE_{it}$	The quarterly change in leverage.
MKT_SHARE_{it}	The percentage of market share by four-digit SIC code based on sales for the quarter.
ΔMKT_SHARE_{it}	The quarterly change in MKT_SHARE.
$OPACITY_{it}$	Discretionary accruals, defined as the residuals from a performance controlled Jones (1991) regression as outlined in Kothari et al. (2005).
$\Delta OPACITY_{it}$	The quarterly change in Opacity.
$NOISE_{it}$	-1 multiplied by the correlation coefficient between daily returns and lagged daily returns for each of the two following periods: (1) the three month period before the filing of the 10-Q and (2) the three month period following the filing of the 10-Q. The month of filing of the 10-Q is excluded from both periods.
$\Delta NOISE_{it}$	The change in NOISE around the filing of the 10-Q.
$QBTM_{it}$	The natural logarithm of book-to-market. $\ln(ceqq/(prccq*cshoq))$.
$\Delta QBTM_{it}$	The quarterly change in QBTM.
$QSIZE_{it}$	$\ln(prccq*cshoq)$.

$\Delta QSIZE_{it}$	The quarterly change in QSIZE.
ROE_{it}	Earnings (ibq) scaled by equity (either seqq, ceqq+pstkq, or atq-ltq, in that order).
ΔROE_{it}	The quarterly change in return on equity.
$TURNOVER_{it}$	vol/shrout as of the end of the quarter taken from CRSP.
$\Delta TURNOVER_{it}$	The quarterly change in TURNOVER.

References

Alti, A., Kaniel, R., & Yoeli, U. (2012). Why do institutional investors chase return trends? *Journal of Financial Intermediation*, 21 (4), 694-721.

Amihud, Y., & Mendelson, H. (1987). Trading mechanisms and stock returns: an empirical investigation. *Journal of Finance*, 42 (3), 533-553.

Baginski, S.P. (1987). Intra-industry information transfers associated with management forecasts of earnings. *Journal of Accounting Research*, 25 (2), 196-216.

Barber, B.M., & Odean, T. (2008). All that glitters: the effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies*, 21 (2), 785-818.

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

Baum, C.F. (2006). An introduction to modern econometrics using Stata. College Station: Stata Press.

Bhojraj, S., & Lee, C.M. (2002). Who is my peer? A valuation-based approach to the selection of comparable firms. *Journal of Accounting Research*, 40 (2), 407-439.

Bhushan, R. (1989). Firm characteristics and analyst following. *Journal of Accounting and Economics*, 11 (2-3), 255-274.

Black, F. (1986). Noise. *Journal of Finance*, 41 (3), 529-543.

Bloomfield, R.J. (2002). The 'incomplete revelation hypothesis' and financial reporting. *Accounting Horizons*, 16 (3), 233-243.

Brennan, M.J., & Hughes, P.J. (1991). Stock prices and the supply of information. *Journal of Finance*, 46 (5), 1665-1691.

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 (Supplement), 171-202.

Bushman, R.M., Piotroski, J.D., & Smith, A.J. (2004). What determines corporate transparency? *Journal of Accounting Research*, 42 (2), 207-252.

Chamley, C. (2007). Complementarities in information acquisition with short-term trades. *Theoretical Economics*, 2 (4), 441-467.

Chan, K., & Hameed, A. (2006). Stock price synchronicity and analyst coverage in emerging markets. *Journal of Financial Economics*, 80 (1), 115-147.

Clinch, G., & Lombardi, B. (2011). Information and the cost of capital: the Easley-O'Hara (2004) model with endogenous information acquisition. *Australian Journal of Management*, 36 (1), 5-14.

Clinch, G.J., & Sinclair, N.A. (1987). Intra-industry information releases. A recursive systems approach. *Journal of Accounting and Economics*, 9 (1), 89-106.

Cohen, L., & Lou, D. (2012). Complicated firms. *Journal of Financial Economics*, 104 (2), 383-400.

Crawford, S.S., Roulstone, D.T., & So, E.C. (2012). Analyst initiations of coverage and stock return synchronicity. *The Accounting Review*, 87 (5), 1527-1553.

Da, Z., Endelberg, J., & Gao, P. (2011). In search of attention. *The Journal of Finance*, 66 (5), 1461-1499.

Damodaran, A. (1993). A simple measure of price adjustment coefficients. *Journal of Finance*, 48 (1), 387-400.

Datta, S., Iskandar-Datta, M. & Singh, V. (2014). Opaque financial reports and R^2: Revisited. *Review of Financial Economics*, 23 (1), 10-17.

De Franco, G., Hope, O., & Larocque, S. (2015). Analysts' choice of peer companies. *Review of Accounting Studies*, 20 (1), 82-109.

Durnev, A., Morck, R., & Yeung, B. (2004). Value-enhancing capital budgeting and firm-specific stock return variation. *Journal of Finance*, 59 (1), 65-105.

Durnev, A., Morck, R., Yeung, B., & Zarowin, P. (2003). Does greater firm-specific return variation mean more or less informed stock pricing? *Journal of Accounting Research*, 41 (5), 797-836.

Ettredge, M., Johnstone, K., Stone, M., & Wang, Q. (2011). The effects of firm size, corporate governance quality, and bad news on disclosure compliance. *Review of Accounting Studies*, 16 (4), 866-889.

Fama, E.F., & French, K.R. (1997). Industry costs of equity. *Journal of Financial Economics*, 43 (2), 153-193.

Financial Accounting Standards Board (FASB). (2008). Exposure Draft for Conceptual Framework for Financial Reporting: The Objective of Financial Reporting and Qualitative Characteristics and Constraints of Decision-Useful Financial Reporting Information. Stamford, CT: FASB. Available from: http://www.fasb.org/draft/ed_conceptual_framework_for_fin_reporting.pdf

Foster, G. (1981). Intra-industry information transfers associated with earnings releases. *Journal of Accounting and Economics*, 3 (3), 201–232.

Froot, K.A., Scharfstein, D.S., & Stein, J.D. (1993). Risk management: coordinating corporate investment and financing policies. *Journal of Finance*, 48 (5), 1629-1658.

Ganguli, J.V., & Yang, L. (2009). Complementarities, multiplicity, and supply information. *Journal of the European Economic Association*, 7 (1), 90-115.

Gleason, C.A., Jenkins, N.T., & Johnson, W.B. (2008). The contagion effects of accounting restatements. *The Accounting Review*, 83 (1), 83-110.

Gu, F., & Wang, W. (2005). Intangible assets, information complexity, and analysts' earnings forecasts. *Journal of Business Finance & Accounting*, 32 (9-10), 1673-1702.

Hameed, A., Morck, R., Shen, J., & Yeung, B. (2015). Information, analysts, and stock return comovement. *The Review of Financial Studies*, 28 (11), 3153-3187.

Han, J.C.Y., & Wild, J.J. (1990). Unexpected earnings and intra-industry information transfers: further evidence. *Journal of Accounting Research*, 28 (1), 211-219.

Han, B., & Yang, L. (2013). Social networks, information acquisition, and asset prices. *Management Science*, 59 (6), 1444-1457.

Hirshleifer, D., Subrahmanyam, A., & Titman, S. (1994). Security analysis and trading patterns when some investors receive information before others. *Journal of Finance*, 49 (5), 1665-1698.

Hoberg, G., & Phillips, G.M. (2012). The stock market, product uniqueness, and comovement of peer firms. *Working paper*, University of Maryland, University of Southern California.

Hou, K. (2007). Industry information diffusion and the lead-lag effect in stock returns. *Review of Financial Studies*, 20 (4), 1113-1138.

Hutton, A.P., Marcus, A.J., & Tehranian, H. (2009). Opaque financial reports, R^2, and crash risk. *Journal of Financial Economics*, 94 (1), 67-86.

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

Jones, J. (1991). Earnings management during import relief investigations. *Journal of Accounting Research*, 29 (2), 193-229.

Kothari, S.P., Leone, A., & Wasley, C. (2005). Performance matched discretionary accrual measures. *Journal of Accounting and Economics*, 39 (1), 163-197.

KPMG and Financial Executives Research Foundation (KPMG). (2011). Disclosure overload and complexity: Hidden in plain sight. Available from: http://www.kpmg.com/us/en/issuesandinsights/articlespublications/pages/disclosure-overload-complexity.aspx

Lang, M.H., & Lundholm, R.J. (1996). Corporate disclosure policy and analyst behavior. *The Accounting Review*, 71 (4), 467-492.

Lang, M.H., Lins, K.V., & Miller, D.P. (2003). ADRS, analysts, and accuracy: does cross listing in the United States improve a firm's information environment and increase market value? *Journal of Accounting Research*, 41 (2), 317-345.

Lee, Y. (2012). The effect of quarterly report readability on information efficiency of stock prices. *Contemporary Accounting Research*, 29 (4), 1137-1170.

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

Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics*, 45 (2-3), 221-247.

Loughran, T., & McDonald, B. (2014). Measuring readability in financial disclosures. *The Journal of Finance*, 69 (4), 1415–1849

Lys, T., & Soo, L.G. (1995). Analysts' forecast precision as a response to competition. *Journal of Accounting, Auditing & Finance*, 10 (4), 751-765.

Manzano, C., & Vives, X. (2011). Public and private learning from prices, strategic substitutability and complementarity, and equilibrium multiplicity. *Journal of Mathematical Economics*, 47 (3), 346-369.

Mele, A., & Sangiorgi, F. (2015). Uncertainty, information acquisition and price swings in asset markets. *Review of Economic Studies*, 82 (4), 1533-1567.

Miller, B.P. (2010). The effects of reporting complexity on small and large investor trading. *The Accounting Review*, 85 (6), 2107-2143.

Morck, R., Yeung, B., & Yu, W. (2000). The information content of stock markets: why do emerging markets have synchronous stock price movements? *Journal of Financial Economics*, 58 (1-2), 215-260.

Muslu, V., Rebello, M., & Xu, Y. (2014). Sell-side analyst research and stock comovement. *Journal of Accounting Research*, 52 (4), 911-954.

Olsen, C., & Dietrich, R. (1985). Vertical information transfers: the association between retailers' sales announcements and suppliers' security returns. *Journal of Accounting Research*, 23 (Supplement), 144-166.

Opp, C. (2015). Intertemporal Information Acquisition and Investment Dynamics. *Working paper,* University of Pennsylvania.

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.

PricewaterhouseCoopers LLP (PWC), (2011). Reducing Complexity: Our Proposal to Address this Challenge. New York, NY: PWC. Available from: http://www.pwc.com/us/en/point-of-view/assets/POV-reducing-complexity-june-2011.pdf

Roll, R. (1988). R^2. *Journal of Finance*, 43 (3), 541-566.

Securities and Exchange Commission (SEC). (2008). Final Report of the Advisory Committee on Improvements to Financial Reporting to the United States Securities and Exchange Commissions. Available from: http://www.sec.gov/about/offices/oca/acifr/acifr-finalreport.pdf

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

Veldkamp, L. (2011). Information choice in macroeconomics and finance. Princeton: Princeton University Press.

Wurgler, J. (2000). Financial markets and the allocation of capital. *Journal of Financial Economics*, 58 (1-2), 187-214.

You, H., & Zhang, X. (2009). Financial reporting complexity and investor underreaction to 10-K information. *Review of Accounting Studies*, 14 (4), 559-586.

Yuan, Y. (2015). Market-wide attention, trading, and stock returns. *Journal of Financial Economics*, 116 (3), 548-564.