9-2018

Leveling the playing field between large and small institutions: Evidence from the SEC's XBRL mandate

Nilabhra BHATTACHARYA
Southern Methodist University

Young Jun CHO
Singapore Management University, yjcho@smu.edu.sg

Jae B. KIM
Lehigh University

DOI: https://doi.org/10.2308/accr-52000

Leveling the Playing Field between Large and Small Institutions: Evidence from the SEC's XBRL Mandate

By: Bhattacharya, Nilabhra, Southern Methodist University

Cho, Young Jun, Singapore Management University

Kim, Jae B, Lehigh University

Published in Accounting Review, Sep 2018, Vol. 93, Issue 5, 51-71.

https://doi.org/10.2308/accr-52000

Submitted version, June 2015

Abstract

We investigate how XBRL adoption affects smaller institutions' access to financial statement information relative to their larger counterparts. We examine three aspects of trading responsiveness: abnormal trading volume, response speed to 10-K information, and decision to trade immediately following the 10-K filing. With regard to all three aspects of trading responsiveness, we find that small institutions' responsiveness to 10-K news increases significantly more relative to the change experienced by large institutions from the pre- to post-XBRL periods. We further document that small institutions' stock picking skills in the 10-K filing period increase more compared to those of large institutions following the regulation. Our results are robust to a battery of falsification and sensitivity tests. Collectively, our results suggest that the informational playing field between small and large institutions has become more even following the SEC's XBRL mandate.

Keywords

XBRL, analysts, institutions, information asymmetry

I. INTRODUCTION

The eXtensible Business Reporting Language (XBRL) is an Internet-based programming language that disseminates 10-K information in machine-readable formats and automates the process of incorporating financial statement information into end users' decision models. The Securities and Exchange Commission (SEC) required companies to file their 10-K reports using XBRL through a three-year phase-in period from 2009 to 2011. The SEC (2009, 67) comments, "We believe analysts, investors, and other market participants will benefit from the enhanced ability of interactive filing to locate and compare financial data included in registration statements." The empirical evidence, however, is sparse regarding the extent to which the XBRL technology fulfills the Commission's intended goals. Early evidence suggests that the information asymmetry between institutional and individual investors has actually widened following the XBRL adoption ([8]). In this study, we investigate the impact of XBRL among institutional investors. Specifically, we examine whether the information asymmetry between larger and smaller institutions has narrowed following the XBRL adoption.[1]

The U.S. XBRL taxonomies and tagging structures established by the SEC are large and complex, and users need significant financial knowledge to understand them.[2] Further, if an item in a company's financial statement cannot be traced to a specific element in the XBRL taxonomy, then the company can create an extension element for that item. At the initial stage of XBRL disclosures, there was little regulatory guidance in building such extension elements. As a result, numerous redundant financial statement elements were created ([20]). Also, the proportion of errors in XBRL data was initially quite high, and one needed significant financial acumen to detect and rectify these errors ([18]; [26]). Finally, users need technical competency to develop software or modify existing software to access XBRL data ([35]). In summary, it appears that XBRL disclosure, at its early stage of development, might put users without a certain level of financial and technical expertise at a relative disadvantage. Indeed, evidence reported in Blankespoor eta al (2014) [8] is consistent with this contention.

This prompts us to explore whether XBRL narrows the information gap between different classes of institutional investors because these users are arguably more sophisticated than retail investors. Big institutions and brokerage houses have large research departments and state-of-the-art technologies to speed up the information acquisition process ([33]). Moreover, several large trading houses already possess in-house technologies that automate the process of feeding financial information directly into their decision models.[3] Research finds that large institutions tend to trade less around 10-K information, possibly because they have access to alternative information sources (e.g., [4]; [32]). Therefore, it is unclear what benefits big institutions reap from the mandate.

In contrast, smaller and boutique institutions are less likely to possess in-house technologies that automate the acquisition of financial data. Smaller trading firms may also be at a relative disadvantage in terms of hand-coding financial information into machine-readable format due to limited manpower. Contemporaneous work suggests that information processing cost has likely decreased following the XBRL adoption ([16]; [30]; [21]). Insofar as smaller institutions clear the minimum threshold of technical and financial competencies, they stand to benefit more from XBRL's automated filing technology. Hence, we investigate whether the information gap between larger and smaller institutions has narrowed as a result of XBRL adoption.

We compare large and small institutions' trading activities in response to 10-K reports filed before and after the XBRL adoption using a proprietary database compiled by Abel Noser Solutions. The database contains 47 million trade orders and the corresponding complete transaction history of 982 institutions since 1999. Several recent studies in the market microstructure literature employ Abel Noser data to investigate various aspects of institutional trading activities (e.g., [22]; [15]; [36]; [2]). Since our goal is to examine institutional responses to 10-K news, we use these data to track institutional trading activities on a daily basis, starting from the day before the 10-K filing date. Institutional holdings data from 13F disclosures and transaction data from the Trade and Quote (TAQ) database cannot accurately track institutional buy/sell trading activities on a day-to-day basis.[4]

There are obviously systematic differences between different types of professional investors, and the concern is whether our results are attributable to these differences between investor classes, as opposed to the XBRL adoption. To mitigate this concern, we employ a generalized difference-in-differences research design by contrasting the trading responses of large versus small institutions to firms' 10-K news before and after the XBRL mandate during our main sample period from 2007 to 2010. We also include firm fixed effects and institution fixed effects in our regression models to control for time-invariant firm- and institution-specific differences. Furthermore, we include year fixed effects to account for staggered XBRL adoption. Finally, we run a number of falsification tests to ensure that our results are attributable to XBRL adoption and not to other concurrent, but unrelated, events. Notably, we repeat our tests around earnings announcements because if the differential trading behaviors are actually driven by XBRL, then the effects are likely to be concentrated around 10-K filing dates, but not around earnings announcements. Thus, null results around earnings announcements can provide additional assurance that the XBRL mandate is the likely catalyst for the observed differences in trading responses between large and small institutions.

We start by investigating institutions' trading responsiveness to 10-K filings before and after XBRL adoption. We examine three aspects of trading responsiveness: (1) abnormal trading volume, (2) response speed to 10-K information, and (3) decision to trade or not trade around 10-K filing dates. With regard to all three measures, we find that small institutions' responsiveness to 10-K news increases significantly more relative to the change experienced by large institutions from the pre- to the post-XBRL periods. Next, we provide evidence that small institutions' ability to pick the right set of stocks during the 10-K filing period increases more from the pre-XBRL to the post-XBRL periods relative to the change observed for large institutions. In addition, we find that small institutions tend to derive greater benefits from XBRL's automated filing formats when financial reports are more complex. All of our results hold after controlling for a series of factors that are shown to influence institutional trading behavior. Our results are also robust to alternative model specifications and a battery of additional sensitivity tests.

Our results have several important implications. First, our evidence suggests that the information asymmetry between larger and smaller institutions has likely narrowed following XBRL adoption. Although our study cannot comment on the impact of XBRL on the informativeness of retail investors, our evidence nonetheless suggests that the regulatory mandate at least partially fulfills the SEC's intended goals. Second, if the narrowing of the information gap between larger and smaller institutions induces greater competition among them in terms of acquisition and processing of information, then this could make the whole sector more efficient. Given the enormous size of institutional capital in the U.S., such development could benefit the overall economy. Third, researchers and regulators quite often focus on the extent of information asymmetry between individuals and institutions. The study's evidence suggests that it may be premature to judge the efficacy of a regulation designed to level the informational playing field solely on the basis of its impact on the information asymmetry between institutional and individual investors.

The rest of the paper proceeds as follows. Section II provides a brief overview of XBRL and presents our research questions. Section III discusses our data and sample. Section IV describes our research design and presents empirical results. Finally, Section V provides concluding remarks.

II. BACKGROUND AND RESEARCH QUESTIONS

Brief Overview of XBRL and Related Regulatory Initiatives

XBRL is an Internet-based programming language that facilitates automated production and consumption of large volumes of business data. XBRL enables information end users to automate the process of incorporating large volumes of financial statement data into their data warehouses and decision models. Companies rely on XBRL taxonomies to prepare their financial statements for interactive filings. Taxonomies are dictionaries that contain standard definitions of financial statement items. The definition explains what each reporting item captures and how it is represented in standard GAAP-based financial statements. The SEC establishes the XBRL taxonomy and makes it available on a website for users to download into their tagging systems. Companies preparing for XBRL filings tag each reporting item in their financial statements with an element from the U.S. GAAP taxonomy that describes the reporting item.[5] If a particular financial statement item cannot be traced to an element in the U.S. GAAP taxonomy, then the company is allowed to create an extension element for that item. A company-specific XBRL file (called an Instance Document) is developed by mapping the company's financial statement line items to the official XBRL taxonomy and the customized extension elements created by the company. Once the Instance Documents are created, preparers make these documents available on the regulator's website (i.e., the SEC's EDGAR) or on their own corporate websites for users to directly download them into their analytical applications. In the first phase of the mandate, the SEC published a foundation taxonomy with more than 15,000 elements or concepts that represent the common practices and disclosure requirements of the U.S. GAAP ([19]). On April 13, 2009, the SEC made the use of XBRL filing mandatory in the U.S. The Commission implemented the XBRL mandate through a three-year phase-in period from 2009 to 2011.

Impact of the XBRL Mandate on Large and Small Institutions

Although the XBRL technology holds promise to simplify and expedite market participants' access to financial information, it seems that a certain minimum threshold of financial and technical expertise is needed to benefit from XBRL data. First, the U.S. GAAP foundation taxonomy itself is large and complex. In addition, if an item in a company's financial statement cannot be traced to a specific element within the standard U.S. GAAP taxonomy, then companies are allowed to build their own extension elements, but the process has not been standardized. Not surprisingly, numerous unnecessary extension elements are created even when semantically equivalent elements already exist in the U.S. GAAP taxonomy. Debreceny et al (2011) [19] analyze extension elements made in a subset of XBRL filings to the SEC between April 2009 and June 2010 and find that 40 percent of these extensions are unnecessary because equivalent elements already exist in the foundation taxonomy. Furthermore, Bovee, Kogan, Nelson, Srivastava, and Vasarhelyi (2005) [10] argue that designing taxonomies for footnote disclosures could be particularly challenging and could contribute to the confusion. Thus, at its early stage, the XBRL taxonomy was complex and confusing, with many redundant labels. Second, the XBRL data quality was initially poor due to an exceptionally high error rate (e.g., Harris and Morsfield 2012[26], 31). Debreceny et all (2010) [18] compare all XBRL filings up to September 2009 with the corresponding published financial statements to assess the error rate in XBRL data. They find that approximately 25 percent of the filings have errors, with an average of seven errors per filing. An end user without sufficient financial expertise may not be able to detect and rectify these errors. Finally, as Blankspoor et all (2014 [8] note, investors have to develop new software or modify existing software to incorporate XBRL data into their analysis, which may not be a trivial task. The above discussion suggests that early stages of XBRL disclosures may only have benefited users with sufficient domain-based experience and expertise, although the accessibility and reliability of XBRL-formatted disclosures have likely improved over the years.

Since professional investors, unlike retail investors, are more likely to clear the minimum bar of technical and financial expertise, we investigate to what extent the XBRL technology levels the playing field between large and small institutions. The SEC (2009) [37] contends that analysts, investors, and other market participants will benefit from the automated format and search-facilitating features of XBRL files. If the mandate benefits market professionals, then it seems intuitive that smaller institutions, with modest resource bases and research supports, stand to benefit more compared to larger trading houses. Extant research and anecdotal evidence supports this intuition. An anecdotal example is that of Credit Suisse First Boston, where portfolio managers and analysts have access to a proprietary database that archives vast amounts of historical information from the construction industry ([33]). In addition, large institutions trade less around 10-K announcements, perhaps because they have various private sources to obtain similar financial information that 10-K reports make public ([4]; [32]). Further, several large trading houses already have in-house technology similar to XBRL. For example, Morgan Stanley's proprietary analytical framework, ModelWare, was designed to transform company data into machine-readable format that allows adjustment, deconstruction, and other modifications to facilitate inter-firm and inter-industry comparisons.[6] Credit Suisse developed its proprietary HOLT framework primarily in response to Morgan Stanley's ModelWare. Consequently, a technology that automates the delivery of 10-K information may be less useful to large institutions.

In contrast, many smaller brokers and institutions do not have access to such sophisticated technological infrastructure, and they could ultimately benefit from XBRL-tagged interactive data. Several small fund managers and analysts from boutique investment shops echoed this sentiment in the survey published by the Chartered Financial Analyst Institute Center for Financial Market Integrity (CFA Institute 2007).[7] Kim et al (2014 [30] report that post-XBRL filings attract more individual and foreign investors, consistent with the notion that XBRL adoption has decreased information processing costs. Again, smaller institutions with limited infrastructure and resources stand to benefit more from such a development.

However, it cannot be known with certainty whether smaller institutions are at a significant disadvantage relative to their larger counterparts in terms of accessing and processing 10-K information. Smaller institutions also gather financial statement information from various alternative sources, so the information contained in the 10-K reports is at least partially preempted prior to the public announcement. Hence, it is not entirely clear whether there was an appreciable information gap between large and small institutions with regard to 10-K information even prior to the regulatory mandate. On the other hand, it is plausible that with their superior infrastructure in place, large institutions may be better placed to exploit XBRL's automated filing format, at least initially, compared to small institutions. If that is the case, then the information asymmetry between them could actually widen, instead of narrowing, during the early phases of XBRL adoption. Consequently, whether the XBRL mandate levels the informational playing field between small and large institutions remains an open empirical question. Hence, the study investigates the following research question:

RQ: Does the informational playing field become more even between smaller institutions and larger institutions following the mandated XBRL adoption?

III. DATA AND SAMPLE SELECTION

The SEC implemented the XBRL mandate through a three-year phase-in period from 2009 to 2011. In the first phase-in period, the SEC required large accelerated filers with a public float of at least $5 billion to file 10-K reports using the XBRL format for fiscal years ending on or after June 15, 2009. In the second phase-in period, public filers with a public float of at least $700 million were required to submit 10-K reports using XBRL-tagged data for fiscal periods ending on or after June 15, 2010. In the final phase-in period, all remaining public filers were mandated to submit XBRL-tagged 10-K reports for fiscal periods ending on or after June 15, 2011. We examine institutional trading responses to corporate 10-K reports filed before and after the XBRL adoption over the period 2007 to 2012. We hand-collect the filing date and the filing format (whether the filing is XBRL-tagged or not) from the SEC's EDGAR website. For each firm in our sample, we label the first year the company filed using the XBRL format as the XBRL adoption year.[8]

To accurately track institutional investors' buy/sell trading activities on a day-to-day basis around corporate 10-K announcements, we use a proprietary database of institutional trades of U.S. equity compiled by Abel Noser Solutions, a consulting firm that helps institutional investors analyze transaction costs. The database uses a unique identifier for each institution and provides the complete transaction history of institutional orders, including the Committee on Uniform Security Identification Procedures (CUSIP) codes of stocks traded, execution date, execution time, execution price, the number of shares executed, an indicator of whether the execution is a buy or a sell, and the commissions, fees, and taxes paid on the execution.[9] We follow Anand et al. (2013) and define a daily trade order (hereafter, order) as the aggregation of all executions (buy and sell) by an institution in the same stock on the same day.

This database has two important advantages over TAQ data. First, by using Abel Noser data, one can precisely track institutional trading activities, while a TAQ user has to rely on noisy trade size-based cutoffs to identify institutional trades. Distinguishing institutional trades from retail trades using transaction size-based cutoffs has become increasingly unreliable. Decimalization in the NYSE and NASDAQ substantially reduced trading costs and prompted institutions to split orders. Also, algorithmic program trading became widespread and further exacerbated institutional order splitting. As a result, the average trade sizes in the NYSE and NASDAQ have decreased sharply from the mid-2000s ([3]). Consequently, using trade size-based proxies to distinguish institutional trades from individual trades has become highly unreliable from the mid-2000s (e.g., [28]; [12]). Second, unlike Abel Noser data, TAQ does not contain information on trade direction, and researchers have used different variations of the Lee and Ready (1991) algorithm to infer buy/sell execution directions. However, the explosive growth in high-frequency trading has significantly reduced the power and accuracy of the Lee-Ready-type algorithms based on TAQ data ([27]).[10]

Abel Noser stopped reporting unique institutional identifiers from 2011 to adequately protect their clients' privacy. Thus, the sample period of our primary analysis is from 2007 to 2010, the period over which we could link each trade order to a specific institution. From 2011, Abel Noser provides unique identifiers of brokers through whom institutions execute their orders. Consequently, in a supplementary analysis, we extend our sample period to December 31, 2012 by using small (large) brokers as proxies for small (large) institutions for the years 2011 and 2012. Large institutions generally trade through large brokers ([7]), so broker size is used as a proxy for institution size. However, it is important to note that this is a noisy proxy and, as a result, our tests based on the extended sample are likely to be less powerful.

We obtain financial statement data from Compustat and stock price/return data from CRSP. The final sample for our institutional trading analysis is obtained by merging the abovementioned databases. Each observation in our institutional trading analysis is at the institution-stock-year level.

Following Anand et al (2013), we impose the following filters on Abel Noser data to minimize reporting errors and eliminate very thinly trading institutions: (1) an order is excluded if its execution price departs from the stock's CRSP opening price by more than 10 percent, (2) an order is eliminated if the order volume is greater than the stock's CRSP volume on the execution date, (3) an order whose size is greater than the 99th percentile of all order sizes in the same month is excluded, and (4) an institution-year-month observation with less than 100 orders is eliminated.[11]

Furthermore, it is very important for this inquiry to identify institutions that pursue an information-based trading strategy and trade vigorously in response to periodic corporate announcements. Prior research (e.g., [11]; [29]) documents that certain institutions have short-term focus, high portfolio turnover, and prompt reactions to corporate announcements (transient institutions), while another group of institutions has long-term focus, follows a passive buy-and-hold strategy, and is characterized by low portfolio turnover (non-transient institutions). Therefore, if the XBRL format does alter institutional trading behavior around 10-K announcements, then the effect is likely to be pronounced in the trading activities of transient institutions, whereas the trading behavior of non-transient institutions should be largely unaltered. Using Abel Noser data, we classify an institution as active/transient based on the number of days the institution trades a stock during a calendar year. This identification strategy is in line with the extant literature because transient (non-transient) institutions are characterized by high (low) portfolio turnovers. We adopt a modest assumption that an institution is active if it trades a stock more than 15 days during a calendar year, and we retain the observation in our sample.[12] We also rerun all of our analyses (untabulated) using alternative cutoffs of 20, 25, and 30 days of trading in a calendar year, and our inferences are unchanged.

In addition, we exclude firms from the sample if they file their 10-Ks more than 120 days after the fiscal year-end, report a negative market-to-book ratio, or have missing values for abnormal trading volume or other control variables. Moreover, if a firm's earnings announcement falls within its own three-day 10-K release window centered on its 10-K announcement date (i.e., days −1 to +1, where day 0 is the 10-K release date), then we eliminate the firm to ensure that we do not inadvertently capture earnings information-induced trading.[13] Finally, we restrict the sample to firms that are traded at least once by both small and large institutions in each calendar year during our sample period. Excluded observations do not seem to reveal strong systematic patterns so as to bias our inferences.

IV. RESEARCH DESIGN AND EMPIRICAL RESULTS

Effect of XBRL on Large and Small Institutions' Trading Responsiveness to 10-K Information

We first probe the impact of the XBRL mandate on the trading behaviors of large and small institutions around firms' 10-K announcement dates. Larger institutions employ a greater number of researchers, analysts, and traders, and maintain large and expensive trading infrastructures ([17]; [33]). Deploying greater resources for trading research and activity is economically viable only if the entity submits large and frequent orders and generates greater trading volume. Consequently, we use the institutional dollar trading volume over an entire year to classify an institution as small or large. Our classification scheme is based on the following approach. Each year, we rank institutions on the basis of their cumulative dollar trading volumes during the entire year. If an institution falls in the top quartile of this distribution, then it is deemed a large institution, while if it is in the bottom three quartiles, it is classified as a small institution. Our inferences are unchanged if we classify large/small institutions using the tercile-based cutoff. In yet additional sensitivity tests, we use the decile rank or the continuous value of the aggregate dollar volume of institutional transactions during the year as alternative measures of institution size, and our results are qualitatively similar. The analyses reported in this section are based on corporate 10-K reports filed between 2007 and 2010.

We examine three different aspects of trading responsiveness: (1) abnormal trading volume, (2) response speed to 10-K information, and (3) decision to trade or not trade surrounding the 10-K release dates. We start by examining small and large institutions' abnormal trading volume during the 10-K announcement period. This measure is computed at the institution-stock-year level. For example, in order to compute Fidelity's abnormal trading volume in response to General Electric's (GE) fiscal 2007 10-K announcement, we only consider the transactions where Fidelity buys or sells GE shares surrounding GE's 2007 10-K announcement period. If Fidelity also trades Microsoft stocks around GE's 10-K release date, then those transactions are ignored for our computation of Fidelity's abnormal volume in response to GE's 10-K filing. Our measure of abnormal trading volume is based on the dollar volume of shares traded and is labeled as AVOL. It is computed as the average daily dollar value of shares of a firm traded by an institution over the three-day window centered on the firm's 10-K filing date (day 0) minus the average daily dollar value of shares of the same firm traded by the same institution over the pre-filing period of days −10 to −2. This measure is then scaled by the average daily dollar value of shares transacted by the same institution-stock pair over the days −10 to +1.[14] If an institution-stock pair has no transaction during the entire −10 to +1 window, then it is eliminated. In untabulated analyses, we repeat all of our tests using an alternative abnormal volume metric based on the number of shares traded instead of the dollar volume of shares traded, and our inferences are unchanged using this alternative measure.[15]

To assess the differential impact of XBRL adoption on large vis-à-vis small institutions' announcement period abnormal volume, we estimate the following model:[16]

$$\begin{aligned} AVOL = {} & \alpha_0 + \alpha_1 XBRL + \alpha_2 SMALL + \alpha_3 (XBRL \times SMALL) + \alpha_4 LOG(MV) + \alpha_5 MTB + \alpha_6 MOM + \alpha_7 WC \\ & + \alpha_8 ABS(10KCAR) + \alpha_9 ABS(EACAR) + \alpha_{10} AFTEAD + \alpha_{11} ONTIME + \alpha_{12} STYLE + \alpha_{13} COST \\ & + \{Institution\ FE\} + \{Firm\ FE\} + \{Year\ FE\} + \varepsilon. \end{aligned} \tag{1}$$

In this model, each observation/record is at the institution-firm-year level. We use all observations from 2007 to 2010, i.e., firms that adopted XBRL by 2010 (adopters), as well as those that did not adopt XBRL by the end of 2010 (non-adopters). XBRL is an indicator variable defined at the firm-year level. XBRL takes the value of 1 if a firm whose stocks are traded by our sample institutions files its 10-K in a given year using XBRL-tagged data, and it is coded 0 if the company files using the traditional HTML format. Since our sample includes both adopters and non-adopters, the coefficient on XBRL facilitates a comparison between pre- and post-XBRL adoption years within adopters after taking into account events affecting overall institutional trading behavior during our sample period. SMALL is an indicator variable intended for capturing the difference between small and large institutions, and it is defined at the institution-year level. SMALL takes the value of 1 if a sample institution in a given year is classified as a small institution, and it is coded 0 otherwise. The variable obtained by interacting XBRL with SMALL (XBRL × SMALL) captures the incremental effect of XBRL adoption on the abnormal trading volume of smaller institutions relative to that of larger institutions. This is our main variable of interest, and a significantly positive coefficient on this interaction term ($\alpha3$) will indicate that smaller institutions' responsiveness to 10-K information has increased more compared to that of larger institutions in the post-XBRL period after controlling for factors affecting overall institutional trading patterns.

Our first set of control variables includes firm attributes that are related to institutional ownership (e.g., [23]). These variables are firm size (LOG(MV)), market-to-book ratio (MTB), and stock momentum (MOM). They are measured at the firm-year level. LOG(MV) is the natural log of the product of the number of shares outstanding and stock price; both are measured at the fiscal year-end of the 10-K report. MTB is the ratio of the market value of common equity to the book value of common equity; again, they are measured at the fiscal year-end of the 10-K report. The stock momentum (MOM) of a company is measured as daily stock returns compounded over a 90-day period ending one day before the firm's 10-K release date minus daily market returns compounded over the same period.

In addition, following prior literature, we control for firm characteristics that influence investors' reaction to 10-K information. We include the natural logarithm of the number of words in a firm's 10-K report (WC) because Miller (2010) documents that longer filings are associated with relatively lower trading. Our ABS(10KCAR) variable is included to control for the magnitude of surprise in the 10-K announcement, and it is measured as the decile ranking of the absolute value of cumulative market-adjusted returns over a three-day period centered on the 10-K filing date.[17] Likewise, we use

ABS(EACAR) to control for the magnitude of surprise in the firm's same fiscal year annual earnings announcement preceding the 10-K filing, and it is measured as the decile ranking of the absolute value of cumulative market-adjusted returns over a three-day period centered on the firm's annual earnings announcement date. We also include a variable that captures the number of calendar days between a firm's earnings announcement date and its 10-K filing date (AFTEAD) to account for cases where earnings information continues to be incorporated during the 10-K release window. Finally, the intensity of market reaction is related to the timeliness of a corporate disclosure ([13]). We include an indicator variable ONTIME that takes the value of 1 if the 10-K is filed within one day from the expected filing date, and it is 0 otherwise. The expected filing date is the same day of the month of last year's 10-K filing.

Furthermore, we control for several additional factors that could affect institutional trading responsiveness. Anand et al (2013) report that institutions have different trading styles in the sense that certain institutions trade more often with the market and serve as long-term liquidity demanders, while other institutions trade more often against the market and serve as long-term liquidity suppliers. Consequently, we control for trading style (STYLE) for our sample institutions. Following an approach similar to that of Anand et al (2013), we calculate STYLE for each institution in our sample based on the percentage of annual trading volume in the same direction as the contemporaneous daily returns of the stocks traded by each institution.[18] The direct transaction costs of trading have been steadily declining over time, and this could disproportionately benefit smaller institutions because they submit relatively smaller orders and, as a result, they face a higher cost per trade compared to larger institutions. To ensure that this temporal trend does not drive our results, we include a proxy for institutions' direct trading costs (COST). COST is calculated as an institution's aggregate trading commissions, fees, and taxes in a year, scaled by the total number of shares traded by the institution in that year.

Finally, we include institution fixed effects to control for omitted institutional attributes that may act as potential confounds. We include firm fixed effects to control for variations in all time-invariant firm-specific attributes. Year fixed effects account for changes over time, since not all firms adopted XBRL in the same year.

The speed of reaction to 10-K information is the second aspect of trading responsiveness we investigate. If XBRL adoption eases the constraint of HTML-formatted disclosures not being readily machine-readable, then small institutions should enjoy speedier access to 10-K information after the regulatory mandate. As a result, the speed of their response to 10-K news should increase more than that of large institutions from the pre- to post-XBRL periods. We compute the response speed to 10-K information as the total dollar volume of shares of a firm traded by an institution during the three-day period centered on the 10-K filing date (days −1 to +1), divided by the total dollar volume of shares of the same firm traded by the same institution over the seven-day period starting from the day before the filing date (days −1 to +5). We label this measure of response speed as SPEED. As mentioned earlier, we compute a second measure of response speed exactly the same way, except using the number of shares traded instead of the dollar volume of shares, and untabulated analyses reveal virtually identical results. These metrics capture the speed of reaction to 10-K news in the following sense. If 10-K information can be accessed and processed faster, then one would expect investor trades to be tightly clustered around the 10-K announcement. On the other hand, if information is obtained and processed more slowly, then investor trades will be dispersed over several days following the 10-K announcement. Thus, our speed measure captures what proportion of total announcement period trades (assuming that the announcement period stretches up to five days after the filing) occurs within one day after the filing date.[19] A higher value of this measure indicates that institutional trades are tightly clustered around 10-K announcements and hence, the response speed is greater. If the XBRL format automates the process of incorporating 10-K information directly into end users' decision models and allows small institutions to process 10-K information faster, then small institutions' response speed should increase more than that of large institutions after the XBRL mandate. We test this conjecture by estimating the following equation:

$$SPEED = \beta_0 + \beta_1 XBRL + \beta_2 SMALL + \beta_3 (XBRL \times SMALL) + \{Controls\} + \{Institution\ FE\} + \{Firm\ FE\} + \{Year\ FE\} + \varepsilon. \quad (2)$$

Again, we use all observations from 2007 to 2010 while estimating Equation (2), i.e., we include both adopters and non-adopters. Again, XBRL is an indicator variable defined at the firm-year level, and it takes the value of 1 if a firm files its 10-K in a given year using the XBRL format, and it is coded 0 if a firm files using the HTML format. Since our sample includes both adopters and non-adopters, the XBRL variable facilitates a comparison between pre- and post-adoption years within adopters. SMALL is defined at the institution-year level and takes the value of 1 if a sample institution in a given year is classified as a small institution, and it is coded 0 otherwise. Finally, the (XBRL × SMALL) interaction term is our main variable of interest, and a significantly positive coefficient on it (β3) will indicate that smaller institutions' reaction speed to 10-K information has increased more compared to that of larger institutions from the pre- to the post-XBRL periods. We include the same set of control variables as in Equation (1) in this regression. Also, as before, the model is estimated using institution fixed effects, firm fixed effects, and year fixed effects.

The third aspect of trading behavior we investigate is the decision to trade or not trade in response to 10-K information. Note that our goal is to examine 10-K information-induced trading and not trading in general. Our abnormal volume test (specified in Equation (1)) can detect whether small institutions' abnormal volume has increased more relative to that of large institutions following the regulatory mandate. Note that abnormal volume could increase if small institutions continue to trade the same set of firms they were trading before, but simply trade more shares of these same firms following 10-K announcements after the XBRL mandate. However, abnormal volume in the 10-K announcement period could also increase if small institutions decide to trade firms in the post-XBRL disclosure regime that they did not trade during these firms' 10-K announcement windows in the pre-XBRL regime. In order to probe this issue, we convert our abnormal volume metric (AVOL) into a dichotomous measure, labeled as TRADING. Thus, TRADING is an indicator variable that takes the value of 1 if AVOL is positive (that is, greater than zero); otherwise, it assumes the value of 0.[20] The average value of this dichotomous trading responsiveness measure will be higher in the post-XBRL period compared to the pre-XBRL period only if an institution expands its 10-K information-induced trading coverage in the post-period relative to the pre-period.[21]

However, interpretation of this metric is problematic in the following scenarios. If an institution does not trade a firm's stocks at all in the pre-XBRL period, but starts trading the firm's stocks during its 10-K release window in the post-XBRL period, then there would be no observation in the pre-XBRL period for that institution-stock pair. Conversely, if an institution trades a firm's stocks in the pre-period, but simply stops covering the firm altogether in the post-period, then there would be no corresponding observation in the post-XBRL period. To ensure that unbalanced pre-/post-panels do not influence our inferences, for this test, we restrict the sample only to institution-stock pairs that have at least one observation in both the pre-XBRL and post-XBRL periods. In essence, our "decision to trade" analysis will only include firms that have adopted XBRL by the end of 2010 (i.e., adopters). Non-adopters will have no observation in the post-XBRL period and will be eliminated. Consequently, this analysis is based on a smaller subsample of observations. We estimate the following model to investigate the differential impact of XBRL adoption on large vis-à-vis small institutions' choice to trade surrounding 10-K releases:

$$\begin{aligned} TRADING = {} & \gamma_0 + \gamma_1 XBRL + \gamma_2 SMALL + \gamma_3 (XBRL \times SMALL) + \{Controls\} + \{Institution\ FE\} + \{Firm\ FE\} \\ & + \{Year\ FE\} + \varepsilon. \end{aligned} \tag{3}$$

The dependent variable in the above equation is our dichotomous "decision to trade" measure. As such, we estimate Equation (3) using a probit model.[22] Note that the right-hand side of this equation is identical to Equations (1) and (2). Again, the main variable of interest is the (XBRL × SMALL) interaction term, and a significantly positive coefficient on this interaction (γ3) will indicate that smaller institutions have expanded their trading coverage more relative to the change observed for large institutions from the pre- to the post-XBRL disclosure regimes.

We estimate Equations (1), (2), and (3) using our data panel spanning 2007 through 2010. Observations of the same firm over time are often correlated. Similarly, there are likely cross-sectional correlations in institutions' trading activities if these institutions are trading during the same time. To purge the effects of these correlations, we cluster our standard errors by firm and by year-month in all of our regression models.[23]

Table 1 provides information on the number of institutions, number of stocks/firms, and number of institution-stock pairs included in each year of our sample period. The table also reports firms and institution-firm pairs that adopted XBRL versus those that did not during each year. For example, Panel A shows that 949 (936 + 13) firms were traded by our sample institutions in 2009, and 843 (557 + 286) firms were traded in 2010. Note that the same firms could appear in our count of 949 in 2009 and 843 in 2010 because the same firms could be traded again next year by the same institution or by other institutions. As expected, very few firms adopted XBRL in 2009, and a significant proportion adopted XBRL in 2010. Also, the table shows that the distribution of observations into adopter and non-adopter groups in our final sample is not lopsided; both groups contain a sufficient number of firms and institution-firm pairs. The table further breaks this information down separately for large and small institutions in Panels B and C, respectively. Note that both Panels B and C contain the same number of firms in each year because, as mentioned in Section III, we require that each firm included in our sample be traded at least once each calendar year by both small and large institutions.

TABLE 1 Yearly Distribution of Sample Institutions and Firms

Panel A: All Institutions

Year	No. of Institutions	No. of Stocks with XBRL = 0	No. of Stocks with XBRL = 1	No. of Institution-Stocks with XBRL = 0	No. of Institution-Stocks with XBRL = 1
2007	229	1,133	0	16,265	0
2008	203	1,064	0	14,306	0
2009	185	936	13	11,648	311
2010	161	557	286	4,398	5,158

Panel B: Large Institutions

Year	No. of Institutions	No. of Stocks with XBRL = 0	No. of Stocks with XBRL = 1	No. of Institution-Stocks with XBRL = 0	No. of Institution-Stocks with XBRL = 1
2007	67	1,133	0	11,762	0
2008	58	1,064	0	9,882	0
2009	55	936	13	8,097	209
2010	48	557	286	3,277	3,678

Panel C: Small Institutions

Year	No. of Institutions	No. of Stocks with XBRL = 0	No. of Stocks with XBRL = 1	No. of Institution-Stocks with XBRL = 0	No. of Institution-Stocks with XBRL = 1
2007	162	1,133	0	4,503	0
2008	145	1,064	0	4,424	0
2009	130	936	13	3,551	102
2010	113	557	286	1,121	1,480

This table provides information on the number of institutions, number of stocks/firms, and number of institution-stock pairs included in each year of our sample period. The table also reports firms and institution-firm pairs that adopted XBRL versus those that did not during each year. Panels B and C contain the same number of firms in each year because we require that each firm included in our sample be traded at least once each calendar year by both small and large institutions.

Table 2 reports descriptive statistics of the dependent and independent variables included in our regression models. Our abnormal trading volume metric (AVOL) has the highest number of observations. We lose more observations when calculating our reaction speed (SPEED) and stock picking measures (defined later). Note that our "decision to trade" measure (TRADING) is based on a much smaller subset of observations because, as mentioned earlier, for this test, we restrict the sample only to institution-stock pairs that have at least one observation in both the pre- and post-XBRL periods. As expected, the abnormal trading volume measure is positive, on average, suggesting above-normal trading activities in the 10-K release window.

TABLE 2 Summary Statistics

Variables	n	Mean	STD	Q1	Median	Q3
AVOL	52,086	0.0327	1.8584	−1.3333	−1.0893	1.0043
SPEED	39,191	0.4814	0.4077	0.0130	0.4506	0.9809
TRADING	9,630	0.3499	0.4770	0.0000	0.0000	1.0000
STOCKPICKING 1	30,264	−0.0001	0.0472	−0.0218	0.0000	0.0223
STOCKPICKING 2	30,264	−0.0006	0.0313	−0.0114	0.0000	0.0109
LOG(MV)	52,086	8.7723	1.7161	7.4743	8.6679	10.0485
MTB	52,086	3.5074	3.3348	1.6237	2.5350	4.0854
MOM	52,086	0.0245	0.1555	−0.0614	0.0167	0.0976
WC	52,086	10.7698	0.4655	10.4760	10.7246	11.0273
ABS(10KCAR)	52,086	0.0262	0.0285	0.0076	0.0173	0.0341
ABS(EACAR)	52,086	0.0535	0.0506	0.0168	0.0374	0.0751
AFTEAD	52,086	22.8623	10.8273	15.0000	23.0000	30.0000
ONTIME	52,086	0.4370	0.4960	0.0000	0.0000	1.0000
STYLE	52,086	0.0494	0.0876	0.0036	0.0516	0.1003
COST	52,086	0.0259	0.0122	0.0179	0.0258	0.0327

This table reports summary statistics for the variables used in our analyses. To mitigate the effects of data errors and outliers, continuous variables are winsorized at the 1 percent and 99 percent levels.
All variables are defined in Appendix A, except that *ABS(10KCAR)* and *ABS(EACAR)* are raw values, not decile rankings.

Table 3 reports the results of estimating Equations (1) through (3), where the three different aspects of trading responsiveness are dependent variables. We find that the coefficients on the interaction term (XBRL × SMALL) are highly significantly positive for all aspects of trading responsiveness—AVOL, SPEED, and TRADING—after controlling for a host of known determinants of institutional trading activities. Furthermore, our difference-in-differences (DiD) research design and inclusion of institution, firm, and year fixed effects in our regression models enable us to isolate the impact of XBRL adoption on the changes in small vis-à-vis large institutions' 10-K information-induced trading responses from the pre- to the post-XBRL disclosure regimes.[24] Consequently, we conclude that we find robust evidence that small institutions' trading responsiveness surrounding 10-K announcement periods has increased significantly more from the pre- to the post-XBRL periods relative to the change experienced by large institutions around 10-K releases.[25] Our results are consistent with the notion that smaller institutions derive greater benefits from the XBRL mandate than larger institutions.

TABLE 3 Small and Large Institutions' Trading Responsiveness to 10-K Information during the Pre- and Post-XBRL Disclosure Regimes

	(1) *AVOL*		(2) *SPEED*		(3) *TRADING*	
Independent Variables	**Coeff.**	**p-value**	**Coeff.**	**p-value**	**Coeff.**	**p-value**
XBRL	−0.1440**	0.017	−0.0007	0.954	−0.0448	0.639
SMALL	−0.0385	0.626	0.0073	0.815	0.1627	0.280
XBRL × *SMALL*	0.1686**	0.018	0.0255**	0.017	0.1855***	0.002
LOG(MV)	−0.0784*	0.096	−0.0252**	0.013	0.0053	0.947
MTB	0.0011	0.923	−0.0005	0.813	0.0172	0.187
MOM	0.0475	0.525	−0.0082	0.623	−0.0060	0.961
WC	0.0122	0.844	0.0023	0.905	0.0888*	0.080
ABS(10KCAR)	0.0116***	0.006	0.0004	0.693	0.0045	0.460
ABS(EACAR)	−0.0053	0.389	−0.0007	0.587	−0.0147*	0.066
AFTEAD	−0.0019	0.485	0.0006	0.457	0.0055	0.209
ONTIME	−0.0595	0.147	−0.0268***	0.007	−0.0082	0.888
STYLE	−0.2129	0.413	0.1214*	0.072	0.0072	0.992
COST	−5.3301	0.369	3.1838***	0.000	−9.9579***	0.001
Institution FE	Included		Included		Included	
Firm FE	Included		Included		Included	
Year FE	Included		Included		Included	
n	52,086		39,191		9,630	
Adjusted (Pseudo) R^2	0.0358		0.0512		0.0540	

***, **, * Denote significance at the 1 percent, 5 percent, and 10 percent levels, respectively.
This table reports the results of examining the differential impact of XBRL adoption on large *vis-à-vis* small institutions' 10-K announcement period trading responsiveness based on corporate 10-K reports filed between 2007 and 2010. The table displays results of estimating the following regression model: $Trading\ Responsiveness = \alpha_0 + \alpha_1 XBRL + \alpha_2 SMALL + \alpha_3(XBRL \times SMALL) + \alpha_4 LOG(MV) + \alpha_5 MTB + \alpha_6 MOM + \alpha_7 WC + \alpha_8 ABS(10KCAR) + \alpha_9 ABS(EACAR) + \alpha_{10} AFTEAD + \alpha_{11} ONTIME + \alpha_{12} STYLE + \alpha_{13} COST + \{Institution\ FE\} + \{Firm\ FE\} + \{Year\ FE\} + \varepsilon$.
The table reports results for three different aspects of trading responsiveness: (1) 10-K announcement period abnormal trading volume (*AVOL*), (2) response speed to 10-K information (*SPEED*), and (3) institutional decision to trade or not trade in the 10-K release window (*TRADING*). The models with *AVOL* and *SPEED* as dependent variables are estimated using all observations—firms that have adopted XBRL by 2010, as well as those that did not adopt XBRL by 2010. The test where *TRADING* is the dependent variable uses a *probit* estimation model and only includes firms that had adopted XBRL by the end of 2010. To mitigate the effects of data errors and outliers, continuous variables are winsorized at the 1 percent and 99 percent levels. The p-values reported in this table are two-sided, and they are based on standard errors clustered by firm and by year-month.
The dependent, independent, and control variables are defined in Appendix A.

In order to further rule out the potential concern that a concurrent event unrelated to the XBRL adoption may be driving the differential trading responses documented thus far, we conduct a series of falsification tests. First, we repeat the same analyses involving AVOL, SPEED, and TRADING around earnings announcements in lieu of 10-K release dates. If the observed differential changes in trading responsiveness are indeed driven by XBRL reporting, then they are more likely to be concentrated around 10-K filing dates and not around annual earnings announcements preceding 10-K filings. [8] run a similar falsification test to substantiate their findings. Hence, we estimate Equations (1) through (3) exactly the same way, except that our three trading responsiveness measures are computed in the short window surrounding annual earnings announcements preceding 10-K filings. The results of these analyses are reported in Table 4. In sharp contrast to the evidence reported in Table 3, Table 4 shows that the coefficients on the (XBRL × SMALL) interaction term are insignificant for all three responsiveness metrics—AVOL, SPEED, and TRADING. These results show that small vis-à-vis large institutions' responsiveness to earnings announcements do not change from the pre- to the post-XBRL disclosure regimes, providing further credence to the notion that the results reported in Table 3 are attributable to the XBRL adoption.

TABLE 4 Falsification Tests Involving Small and Large Institutions' Trading Responsiveness to Annual Earnings Announcements Preceding 10-K filings during the Pre- and Post-XBRL Disclosure Regimes

	(1) *AVOL*		(2) *SPEED*		(3) *TRADING*	
Independent Variables	**Coeff.**	**p-value**	**Coeff.**	**p-value**	**Coeff.**	**p-value**
XBRL	−0.0220	0.694	0.0038	0.793	0.0603	0.660
SMALL	0.2050*	0.098	0.0198	0.165	0.2036	0.229
XBRL × SMALL	−0.0836	0.174	−0.0152	0.170	−0.0163	0.789
LOG(MV)	0.0813**	0.050	0.0105	0.292	−0.0201	0.719
MTB	0.0167*	0.071	0.0037*	0.082	0.0322**	0.011
MOM	0.1192	0.145	−0.0030	0.876	0.1213	0.375
WC	0.0081	0.855	0.0206	0.106	0.0362	0.605
ABS(10KCAR)	−0.0021	0.695	−0.0015	0.165	−0.0079	0.309
ABS(EACAR)	0.0103	0.201	−0.0030**	0.027	0.0058	0.552
AFTEAD	−0.0072***	0.007	−0.0005	0.291	−0.0018	0.720
ONTIME	−0.0084	0.743	0.0024	0.780	0.0028	0.938
STYLE	0.2020	0.556	−0.0100	0.828	0.1296	0.791
COST	1.5993	0.733	1.1769	0.142	4.6547***	0.002
Institution FE	Included		Included		Included	
Firm FE	Included		Included		Included	
Year FE	Included		Included		Included	
n	51,871		40,577		9,738	
Adjusted (Pseudo) R^2	0.0373		0.0498		0.0481	

***, **, * Denote significance at the 1 percent, 5 percent, and 10 percent levels, respectively.

This table reports the results of examining the differential impact of XBRL adoption on large *vis-à-vis* small institutions' trading responsiveness around annual earnings announcements preceding the 10-K filings between 2007 and 2010. These analyses serve as falsification tests. No differential trading responses around earnings announcements reinforce the conclusion that the observed differences between large and small institutions' trading responsiveness are attributable to the XBRL adoption. The table displays the results of estimating the following regression model: $Trading\ Responsiveness = \alpha_0 + \alpha_1 XBRL + \alpha_2 SMALL + \alpha_3(XBRL \times SMALL) + \{Controls\} + \{Institution\ FE\} + \{Firm\ FE\} + \{Year\ FE\} + \varepsilon$.

The table reports results for three different aspects of trading responsiveness: (1) 10-K announcement period abnormal trading volume (*AVOL*), (2) response speed to 10-K information (*SPEED*), and (3) institutional decision to trade or not trade in the 10-K release window (*TRADING*). The models with *AVOL* and *SPEED* as dependent variables are estimated using all observations—firms that have adopted XBRL by 2010, as well as those that did not adopt XBRL by 2010. The test where *TRADING* is the dependent variable uses a *probit* estimation model and only includes firms that had adopted XBRL by the end of 2010. The same set of control variables used in Table 3 is employed here. To mitigate the effects of data errors and outliers, continuous variables are winsorized at the 1 percent and 99 percent levels. The p-values reported in this table are two-sided, and they are based on standard errors clustered by firm and by year-month.

The dependent, independent, and control variables are defined in Appendix A.

Next, we perform a number of additional placebo tests based on the intuition that if the XBRL mandate does alter institutional trading behaviors around 10-K announcements, then the effect is likely to be concentrated in the trading activities of transient institutions, whereas the trading patterns of passive institutions should be largely unaltered because they are unlikely to actively trade around periodic 10-K releases in the first place. In fact, if we observe that non-transient institutions react in a way similar to transient institutions to XBRL adoption, then it may suggest that our results are driven by unrelated events or by some mechanical or spurious associations. In that vein, differential responses from these two groups to the XBRL mandate could serve as additional falsification tests. Therefore, we perform several untabulated placebo tests along this line. As mentioned before, our main tabulated analyses are based on transient institutions (and we delete passive institutions from our sample), relying on the assumption that if an institution trades a stock more than 15 days during a year, it is a transient institution. Analogously, we identify two alternative subsamples of passive institutions: (1) those having less than 15 days of trading in a year, and (2) those having less than ten days of trading in a year. For both subsamples of passive institutions, we find that the coefficients on the (XBRL × SMALL) interaction term are never significantly positive for any of our trading responsiveness metrics (AVOL, SPEED, and TRADING), suggesting that XBRL adoption does not seem to have any effect on non-transient institutions' trading behaviors.[26]

In summary, these series of falsification tests help to convincingly rule out possibilities that some contemporaneous events unrelated to the XBRL adoption are driving our results. Collectively, the numerous analyses we undertake in this section indicate that small institutions trade more vigorously, more swiftly, and expand their trading coverages more during 10-K release windows (i.e., trade stocks that they were not trading earlier in response to the 10-K news) compared to large institutions following the mandated XBRL adoption.

Effect of XBRL on Small and Large Institutions' Stock Picking Abilities in the 10-K Announcement Window

Given the evidence that the machine-readable XBRL format spurs small institutions to trade more, as well as trade promptly, in response to 10-K news, we next investigate whether small institutions' ability to pick the right set of stocks during the 10-K announcement window increases more relative to that of large institutions following the mandate. Since the Abel Noser database contains detailed information on institutional orders and transactions, it enables closer examination of the various aspects of trading responsiveness. However, it is difficult to compute a measure of trading profitability or performance from these data because we have no way of knowing when shares are sold, which batch of prior purchases they come from. Thus, any trading performance measure constructed from these data is going to be noisier than the trading responsiveness metrics we have examined so far. Nevertheless, in this section, we attempt to evaluate trading performance using metrics similar to those used in the mutual fund literature to ascertain the stock picking skills of fund managers ([5]; [7]). These measures determine fund managers' stock picking skills by linking the changes in a fund's stock holdings to subsequent returns of the same stocks.[27] Following the same terminology, we label our trading performance measures as stock picking skills of institutions.

In our stock picking measure, in line with prior work, we link each transaction made by an institution during a firm's 10-K window (i.e., days −1 to +1, where day 0 is the 10-K filing date) to the subsequent return earned from that transaction using day +3 as the last day of the return window. Thus, the return in our stock picking measure is calculated based on the actual transaction price and the closing price on day +3. We then multiply the return with the corresponding dollar value of shares traded in the transaction and an indicator of +1 for purchase or −1 for sale. We compute this product for each transaction involving the institution-stock pair in the event window and sum them together. Finally, this sum is scaled by the total dollar volume of trades from all the transactions in the event window. Thus, the following formula represents our first stock picking measure assuming n number of transactions in the event window:

$$STOCKPICKING\ 1 \equiv \sum_{i=1}^{n} [\{(P_3 - P_i)/P_i\} \times \{P_i \times V_i \times T_i\}] \div \sum_{i=1}^{n} \{P_i \times V_i\},$$

where:

$P_3 \equiv$ closing price at the end of day +3;

$P_i \equiv$ execution price of the *i*th transaction in the event window;

$V_i \equiv$ number of shares traded in the *i*th transaction in the event window; and

$T_i \equiv +1$ if the *i*th transaction is a buy or −1 if the *i*th transaction is a sell.

The above stock picking metric is analogous to a weighted average return measure where the dollar volume of trades tied to each transaction acts as the weights.[28] By construction, this metric will yield a higher value if an institution buys (sells) stocks that are about to experience price upticks (downticks) soon after the 10-K announcement. Consequently, a higher value of this metric indicates better stock picking ability. We define a second stock picking measure, labeled as STOCKPICKING 2, which is a slight variant of our first measure, to show that our results are not sensitive to the cutoff dates for computing transaction-specific returns. STOCKPICKING 2 is calculated exactly the same way, except that the last day of the return cumulation window is day +1.

We estimate the following model to assess the differential impact of XBRL adoption on the changes in large vis-à-vis small institutions' stock picking abilities in the 10-K announcement window:

$$\begin{aligned} STOCKPICKING\ 1 \text{ or } STOCKPICKING\ 2 = {} & \delta_0 + \delta_1 XBRL + \delta_2 SMALL + \delta_3 (XBRL \times SMALL) + \{Controls\} \\ & + \{Institution\ FE\} + \{Firm\ FE\} + \{Year\ FE\} + \varepsilon. \end{aligned} \tag{4}$$

Each record in this model is again measured at the institution-firm-year level. The independent variables are defined earlier. Again, a significantly positive δ3 will indicate that smaller institutions' stock picking skills during the 10-K release period have increased more from the pre- to the post-XBRL regimes compared to that of larger institutions. We include the same set of control variables, institution fixed effects, firm fixed effects, and year fixed effects as in the previous three equations. Again, we cluster standard errors by firm and by year-month while estimating Equation (4).

The results of estimating Equation (4) are reported in Table 5. The (XBRL × SMALL) interaction term is significantly positive for both measures of stock picking skills. These results indicate that small institutions' trading performance, as measured by their stock picking abilities, increases more from the pre- to post-XBRL periods compared to the change experienced by large institutions. The evidence reported in this table is consistent with the inferences derived from our trading responsiveness tests. Collectively, our analyses so far reveal that the informational playing field has become more even between smaller institutions and larger institutions following the mandated XBRL adoption.

TABLE 5 Small and Large Institutions' Stock Picking Skills in the 10-K Filing Period during the Pre- and Post-XBRL Disclosure Regimes

	Regimes			
	(1) *STOCKPICKING 1*		(2) *STOCKPICKING 2*	
Independent Variables	**Coeff.**	**p-value**	**Coeff.**	**p-value**
XBRL	−0.0022*	0.075	−0.0019*	0.072
SMALL	−0.0039	0.390	−0.0026	0.309
XBRL × *SMALL*	0.0029***	0.005	0.0023**	0.020
LOG(MV)	0.0023	0.188	0.0020**	0.037
MTB	−0.0001	0.658	−0.0001	0.258
MOM	0.0026	0.547	0.0066**	0.012
WC	0.0003	0.846	−0.0005	0.605
ABS(10KCAR)	0.0001	0.450	0.0000	0.971
ABS(EACAR)	0.0004***	0.005	0.0002**	0.010
AFTEAD	0.0000	0.899	−0.0001	0.295
ONTIME	0.0006	0.514	−0.0001	0.855
STYLE	0.0113	0.494	0.0194	0.203
COST	−0.6007***	0.004	−0.4016**	0.041
Institution FE	Included		Included	
Firm FE	Included		Included	
Year FE	Included		Included	
n	30,264		30,264	
Adjusted R^2	0.0582		0.0565	

***, **, * Denote significance at the 1 percent, 5 percent, and 10 percent levels, respectively.
This tables reports the results of evaluating the differential impact of XBRL adoption on large *vis-à-vis* small institutions' stock picking skills in the short window surrounding the 10-K announcement for 10-K reports filed between 2007 and 2010. The table reports the results of estimating the following regression model: $STOCKPICKING\ 1$ or $STOCKPICKING\ 2 = \delta_0 + \delta_1 XBRL + \delta_2 SMALL + \delta_3 (XBRL \times SMALL) + \{Controls\} + \{Institution\ FE\} + \{Firm\ FE\} + \{Year\ FE\} + \varepsilon$.
The dependent variable is two slightly different measures of institutions' stock picking skills in the 10-K release window—*STOCKPICKING 1* and *STOCKPICKING 2*. The models are estimated using all observations—firms that have adopted XBRL by 2010, as well as those that did not adopt XBRL by 2010. The same set of control variables used in Tables 3 and 4 is employed here. To mitigate the effects of data errors and outliers, continuous variables are winsorized at the 1 percent and 99 percent levels. The p-values reported in this table are two-sided, and they are based on standard errors clustered by firm and by year-month.
The dependent, independent, and control variables are defined in Appendix A.

Role of XBRL in Facilitating the Assimilation of Complex Financial Reports and Analyses Based...

In this section, we investigate to what extent XBRL's automated filing format facilitates the processing of relatively more complex financial reports. This inquiry is motivated by the argument that information complexity affects the extent of information assimilation and could be one potential source of information asymmetry (e.g., [9]). Extant empirical evidence supports the notion that reporting complexity impedes the extent of information incorporation (e.g., [38]; [32]). Prior to the XBRL mandate, larger institutions likely enjoyed an edge over smaller institutions in processing longer and more complex financial reports due to their superior technological infrastructure and other resources. Consequently, we examine next whether the XBRL adoption has leveled the playing field to a greater extent for relatively more complex disclosures.

Our primary measure of reporting complexity is the total number of tags associated with all financial statement items in an entire XBRL-formatted 10-K filing. In an XBRL filing, each financial statement item (financial statement line items, as well as footnotes) is tagged to entities that provide information regarding the definition of the item, reference to the GAAP standard, reporting period, reporting currency, unit of measurement, and so on ([19]). Complex financial reports are likely to have a greater number of financial statement line items, as well as a greater number of footnotes. These footnotes are also expected to contain a more extensive description of numerous assumptions and complicated measurement processes. Consequently, more complex 10-K filings are likely to be accompanied by a greater number of XBRL tags. One potential concern with this proxy is that the value of complexity is undefined for firms that did not adopt XBRL during our main sample period. To circumvent this problem, we use the XBRL tag count in 2012 (the last year of our extended sample) as a measure of complexity throughout our sample period. Since, in the final phase-in period, all remaining public filers were mandated to submit XBRL-tagged 10-K reports for fiscal periods ending on or after June 15, 2011, the use of XBRL tag count in 2012 avoids the problem of the complexity measure being undefined for non-adopters.[29] In an untabulated analysis, we use file size as an alternative proxy for complexity and find similar results. We estimate the following model to examine how the impact of XBRL adoption on institutional trading reactions in the 10-K filing window is moderated by the complexity of the 10-K reports:

$$\begin{aligned} \textit{Institutional Response} = {} & \lambda_0 + \lambda_1 XBRL + \lambda_2 SMALL + \lambda_3 (XBRL \times SMALL) + \lambda_4 (COMPLEX \times XBRL) \\ & + \lambda_5 (COMPLEX \times SMALL) + \lambda_6 (COMPLEX \times XBRL \times SMALL) + \{Controls\} \\ & + \{Institution\ FE\} + \{Firm\ FE\} + \{Year\ FE\} + \varepsilon. \end{aligned} \tag{5}$$

We estimate the model over our main sample period from 2007 to 2010 and again retain the firms that did not adopt XBRL by 2010 as the benchmark group. The various institutional responses in the 10-K announcement window we have examined so far (i.e., AVOL, SPEED, TRADING, STOCKPICKING 1, and STOCKPICKING 2) act as the dependent variable, one at a time. COMPLEX takes the value of 1 if the number of tags in a firm's 2012 10-K filing is above the top quintile of our sample, and it takes the value of 0 if the tag count falls below the top quintile.[30] Hence, the COMPLEX indicator variable is defined at the firm level. All other independent and control variables are defined earlier. Again, we include institution, firm, and year fixed effects, and cluster standard errors by firm and by year-month when estimating Equation (5). Note that COMPLEX does not enter the above equation as a stand-alone variable because it is a firm-specific variable and, thus, is subsumed by firm fixed effects. In this setting, our focus is on the three-way interaction term (COMPLEX × XBRL × SMALL). Thus, a significantly positive λ6 will indicate that the information asymmetry between large and small institutions narrows to a greater extent when the 10-K filings are relatively more complex.

The results of estimating Equation (5) are reported in Table 6. We find that the three-way interaction term is significant for AVOL, SPEED, and STOCKPICKING 1, while it is insignificant for TRADING and STOCKPICKING 2. We, thus, find modest evidence that the information gap between large and small institutions has narrowed to a greater extent following the regulatory mandate for more complex 10-K filings.

TABLE 6 The Moderating Role of 10-K Reporting Complexity on Small vis-à-vis Large Institutions' Differential Responses to 10-K News during the Pre- and Post-XBRL Disclosure Regimes

	(1) *AVOL*		(2) *SPEED*		(3) *TRADING*		(4) *STOCKPICKING 1*		(5) *STOCKPICKING 2*	
Independent Variables	**Coeff.**	**p-value**	**Coeff.**	**p-value**	**Coeff.**	**p-value**	**Coeff.**	**p-value**	**Coeff.**	**p-value**
XBRL	−0.1593**	0.015	−0.0013	0.934	−0.0697	0.493	−0.0018	0.228	−0.0015	0.171
SMALL	−0.0417	0.593	0.0021	0.940	0.1621	0.276	−0.0036	0.308	−0.0020	0.351
XBRL × *SMALL*	0.1051	0.104	0.0089	0.534	0.1638**	0.026	0.0008	0.638	0.0015	0.253
COMPLEX × *XBRL*	0.0464	0.448	0.0080	0.604	0.1405***	0.009	−0.0012	0.387	−0.0010	0.264
COMPLEX × *SMALL*	−0.0519	0.282	−0.0048	0.761	−0.0071	0.907	−0.0076*	0.062	−0.0048*	0.076
COMPLEX × *XBRL* × *SMALL*	0.2963**	0.017	0.0602**	0.021	0.0742	0.475	0.0096**	0.048	0.0048	0.108
LOG(MV)	−0.0758	0.118	−0.0203*	0.058	0.0079	0.921	0.0022	0.195	0.0017*	0.070
MTB	0.0020	0.857	−0.0008	0.770	0.0143	0.294	−0.0002	0.513	−0.0002	0.101
MOM	0.0382	0.626	−0.0066	0.714	−0.0473	0.685	0.0030	0.536	0.0068***	0.009
WC	−0.0114	0.864	−0.0043	0.827	0.0709	0.165	0.0013	0.407	0.0002	0.776
ABS(10KCAR)	0.0109**	0.013	0.0006	0.604	0.0041	0.500	0.0001	0.480	0.0000	0.666
ABS(EACAR)	−0.0015	0.824	−0.0004	0.758	−0.0147*	0.069	0.0004**	0.010	0.0002**	0.010
EADAFT	−0.0017	0.589	0.0003	0.706	0.0048	0.283	0.0000	0.958	−0.0001	0.167
ONTIME	−0.0536	0.192	−0.0256***	0.010	−0.0032	0.953	0.0002	0.869	−0.0002	0.752
STYLE	−0.1739	0.529	0.1201*	0.056	0.0213	0.976	0.0184	0.296	0.0224	0.170
COST	−4.8022	0.421	3.6286***	0.000	−9.8435***	0.001	−0.6250***	0.003	−0.4138**	0.034
Institution FE	Included		Included		Included		Included		Included	
Firm FE	Included		Included		Included		Included		Included	
Year FE	Included		Included		Included		Included		Included	
n	47,213		35,601		9,534		27,485		27,485	
Adjusted (Pseudo) R^2	0.0336		0.0524		0.0545		0.0564		0.0563	

***, **, * Denote significance at the 1 percent, 5 percent, and 10 percent levels, respectively.

This table examines how the impact of XBRL adoption on the information asymmetry between large and small institutions is moderated by the complexity of the 10-K reports. The table displays the results of estimating the following model: $Institutional\ Response = \lambda_0 + \lambda_1 XBRL + \lambda_2 SMALL + \lambda_3 (XBRL \times SMALL) + \lambda_4 (COMPLEX \times XBRL) + \lambda_5 (COMPLEX \times SMALL) + \lambda_6 (COMPLEX \times XBRL \times SMALL) + \{Controls\} + \{Institution\ FE\} + \{Firm\ FE\} + \{Year\ FE\} + \varepsilon$.

The institutional responses in the 10-K announcement window examined in the previous tables (*AVOL*, *SPEED*, *TRADING*, *STOCKPICKING 1*, and *STOCKPICKING 2*) act as the dependent variable, one at a time. The models are estimated using all observations—firms that have adopted XBRL by 2010, as well as those that did not adopt XBRL by 2010—except when *TRADING* is the dependent variable. The test where *TRADING* is the dependent variable uses a *probit* estimation model and only includes firms that had adopted XBRL by the end of 2010. *COMPLEX* is an indicator variable that measures the complexity of 10-K reports. It takes the value of 1 if the number of tags in the firm's 10-K filed in 2012 is above the top quintile of our sample, and it is 0 otherwise. The same set of control variables used in Tables 3 through 5 is employed here. To mitigate the effects of data errors and outliers, continuous variables are winsorized at the 1 percent and 99 percent levels. The p-values reported in this table are two-sided, and they are based on standard errors clustered by firm and by year-month.

The dependent, independent, and control variables are defined in Appendix A.

In our final set of analyses, we report results based on an extended sample period that covers corporate 10-K reports from 2007 to the end of 2012. Our main analyses reported thus far are based on a sample from 2007 through 2010 because we could link every transaction to a specific institution over this period. Abel Noser stopped reporting unique institutional identifiers from 2011 onward. Instead, they now report identifiers of brokers through whom institutions route their orders. We extend our sample period by using small (large) brokers as proxies for small (large) institutions for the years 2011 and 2012. Large institutions tend to trade through large brokers ([7]), so broker size is used as a noisy proxy for institution size. Note that for the years 2007 through 2010, we still link every transaction to a specific institution. For the years 2011 and 2012, we first rank brokers on the basis of their annual cumulative dollar trading volumes. Each year, if a broker falls in the top quartile of this distribution, it is deemed as a large broker, while if it is in the bottom three quartiles, it is classified as a small broker. Since broker size is an indirect and noisy surrogate for institution size, tests based on our extended sample are probably less powerful than those based on our primary sample.

We reestimate Equations (1) through (4) over this extended sample period, and Table 7 reports these results.[31] Analogous to Panel A of Table 1, Panel A of Table 7 provides information on the distributions of brokers and broker-stock pairs for 2011 and 2012. The panel then breaks down the information further for large and small brokers. The table shows that there are sufficient numbers of large and small brokers in both years. Also, there are ample broker-stock observations in both the adopter and non-adopter groups. Panel B reports the regression results. This panel shows that the coefficient on the (XBRL × SMALL) interaction term is significantly positive for all of our institutional response measures—AVOL, SPEED, TRADING, STOCKPICKING 1, and STOCKPICKING 2. Overall, the results over the extended sample period confirm our earlier findings that small institutions' responses to 10-K news increase significantly more relative to the change observed for large institutions from the pre- to the post-XBRL disclosure regimes.

Graph: TABLE 7 Small and Large Institutions' Responses to 10-K News during the Pre- and Post-XBRL Disclosure Regimes Based on an Extended Sample

Panel A: Yearly Distribution of Sample Brokers

Year	No. of Brokers	No. of Broker-Stocks	No. of Broker-Stocks with XBRL = 0	No. of Broker-Stocks with XBRL = 1	No. of Stocks with XBRL = 0	No. of Stocks with XBRL = 1
All Brokers						
2011	179	8,744	1,412	7,332	211	595
2012	203	6,287	2	6,285	1	756
Large Brokers						
2011	53	6,604	1,096	5,508	211	595
2012	61	4,569	1	4,568	1	756
Small Brokers						
2011	126	2,140	316	1,824	211	595
2012	142	1,718	1	1,717	1	756

Panel B: Regression Results for Trading Responsiveness and Stock Picking Skills

Independent Variables	(1) *AVOL* Coeff.	p-value	(2) *SPEED* Coeff.	p-value	(3) *TRADING* Coeff.	p-value	(4) *STOCKPICKING 1* Coeff.	p-value	(5) *STOCKPICKING 2* Coeff.	p-value
XBRL	−0.1282**	0.017	−0.0058	0.623	−0.0561	0.557	−0.0007	0.526	−0.0006	0.538
SMALL	−0.0066	0.923	0.0077	0.766	0.1629	0.278	−0.0027	0.441	−0.0014	0.480
XBRL × SMALL	0.2047***	0.000	0.0308***	0.003	0.1845***	0.002	0.0021**	0.031	0.0015**	0.039
LOG(MV)	−0.0634	0.105	−0.0308***	0.000	0.0039	0.961	0.0014	0.130	0.0013**	0.039
MTB	0.0033	0.714	−0.0002	0.884	0.0184	0.161	−0.0001	0.561	−0.0001	0.188
MOM	−0.0002	0.998	−0.0046	0.737	−0.0111	0.929	0.0012	0.628	0.0036**	0.046
WC	−0.0446	0.325	0.0095	0.367	0.0761	0.137	0.0002	0.838	0.0002	0.719
ABS(10KCAR)	0.0116***	0.003	−0.0003	0.714	0.0027	0.663	0.0001	0.445	0.0000	0.683
ABS(EACAR)	−0.0038	0.406	−0.0001	0.879	−0.0153*	0.058	0.0003**	0.030	0.0001**	0.039
AFTEAD	−0.0016	0.438	0.0003	0.656	0.0058	0.181	0.0000	0.993	0.0000	0.713
ONTIME	−0.0410	0.257	−0.0196**	0.023	−0.0081	0.887	0.0002	0.794	−0.0001	0.840
STYLE	−0.0615	0.809	0.1154**	0.037	−0.0151	0.983	0.0068	0.591	0.0144	0.200
COST	−2.7435	0.646	3.0045***	0.000	−9.9884***	0.001	−0.4792***	0.006	−0.3361**	0.034
Institution FE	Included		Included		Included		Included		Included	
Firm FE	Included		Included		Included		Included		Included	
Year FE	Included		Included		Included		Included		Included	
n	67,117		54,222		9,784		45,282		45,295	
Adjusted (Pseudo) R^2	0.2914		0.1806		0.0557		0.0526		0.0487	

***, **, * Denote significance at the 1 percent, 5 percent, and 10 percent levels, respectively.

This table reports the results of institutional responses to 10-K news over an extended sample period covering corporate 10-K reports from 2007 to the end of 2012. We extend our sample period by using small (large) brokers as proxies for small (large) institutions for the years 2011 and 2012. For the years 2007 through 2010, we still classify an institution as small or large using its own annual dollar trading volume. Panel A reports the yearly distribution of our sample brokers. Panel B shows regression results for our trading responsiveness and stock picking skill measures using the same set of independent and control variables as in the previous tables. The continuous variables are winsorized at the 1 percent and 99 percent levels. The p-values reported in this table are two-sided, and they are based on standard errors clustered by firm and by year-month.

The dependent, independent, and control variables are defined in Appendix A.

Taken together, the results reported in Tables 3 through 7 provide robust evidence that the information asymmetry between large and small institutions has narrowed following the SEC's XBRL mandate. Therefore, the regulatory mandate seems to at least partially fulfill the Commission's goals.

V. CONCLUSION

In this study, we examine whether the asymmetry between small and large institutions in terms of access to financial statement information has narrowed following the XBRL adoption. As far as the speed of information acquisition and processing is concerned, smaller institutions may have been at a relative disadvantage compared to their larger counterparts. Several large trading houses already possess in-house technologies similar to XBRL that automate the data acquisition process. In contrast, smaller institutions with modest technological infrastructure and limited resources stand to benefit more from XBRL's automated filing formats. Hence, the study investigates whether the informational playing field has become more even between smaller and larger institutions as a result of the XBRL mandate.

We compare large and small institutions' trading activities in response to 10-K reports filed before and after the XBRL adoption using a proprietary database compiled by Abel Noser Solutions over the period 2007 to 2012. We employ these data because the institutional holdings data from 13F disclosures and the transaction data from TAQ are unable to accurately track institutional trading activities on a day-to-day basis surrounding 10-K announcements. We use the institutional annual dollar trading volume to classify an institution as small or large. We report that small institutions have experienced a greater increase in their trading responsiveness to 10-K news following the XBRL mandate compared to large institutions after controlling for a comprehensive set of known determinants of institutional trading activities. We also find that small institutions' stock picking skills in the 10-K announcement window improve more from the pre- to the post-XBRL periods compared to the changes experienced by large institutions. In addition, we find some evidence that small institutions tend to derive greater benefits from XBRL's automated filing formats when financial reports are more complex. Our results are robust to alternative model specifications and a battery of sensitivity analyses and falsification tests.

The SEC's goal of mandating XBRL is to level the informational playing field among various market participants. However, empirical evidence on the extent to which XBRL enhances users' ability to acquire and process information is sparse. Our evidence suggests that the informational playing field between large and small institutions has become more even following the regulatory mandate. In that vein, our study has the potential to inform policymakers and regulators.

We thank two anonymous reviewers, Rachel M. Hayes (editor), Linda Bamber, Xia Chen, Qiang Cheng, Zhaoyang Gu, Santhosh Ramalingegowda, Kumar Venkataraman, Jeff Yu, Yoonseok Zang, workshop participants at The Chinese University of Hong Kong and Singapore Management University, and participants at the 2014 American Accounting Association Annual Meeting for many helpful suggestions. We also thank Shivika Srimal for excellent research assistance. We are grateful to Abel Noser Solutions for providing institutional trading data. We are grateful for the financial support provided by the Cox School of Business at Southern Methodist University and by the School of Accountancy Research Center at Singapore Management University. Young Jun Cho acknowledges funding from the Lee Kong Chian Fellowship.

Notes

1. We classify an institution as small or large using its total annual trading volume. Section IV discusses our classification scheme in detail.
2. Taxonomies are "dictionaries" that contain standard definitions of financial statement items, such as revenue or cash. Companies use XBRL taxonomies to prepare their financial statements for interactive filings. In the first phase of the XBRL mandate, there were more than 15,000 elements in the standard XBRL taxonomy that represented the disclosure requirements of the U.S. generally accepted accounting principles (GAAP), and the list is growing ([19]).
3. For example, Morgan Stanley's proprietary analytical framework, ModelWare, was designed to transform financial statement data into a machine-readable format. In Section II, we provide additional examples of such targeted technologies developed by big financial institutions to automate the data acquisition process.
4. 13F disclosures provide institutional holdings information on a quarterly basis. As such, it is difficult to attribute the changes to 10-K filings, because myriad events over the course of a quarter could influence institutional holdings. Likewise, the TAQ data cannot be used to accurately track institutional trading, because the data do not contain information on trade orders (i.e., whether an order originates from an individual account or an institutional account) or on buy/sell execution directions. The explosive growth in high-frequency trading since the mid-2000s has led to huge increases in execution speed and extensive order splitting by institutions. As a result, the proxies and algorithms employed earlier to distinguish between individual and institutional trading activities and to infer trade execution directions from TAQ are no longer reliable ([28]; [12]; [6]). Section III discusses our data and sample in greater detail.
5. We follow the general convention and use the term "elements" to refer to parts of XBRL taxonomies, while "reporting items" or "items" refer to parts of corporate financial statements ([35]).
6. Morgan Stanley describes ModelWare as a framework that transforms company data into meaningful metrics so that these data are comparable across regions and sectors (see: http://www.morganstanleyiq.ch/EN/binaer%5fview.asp?BinaerNr=141). The company mentions that ModelWare is the result of many years of development by analysts and experts. It operates as an independent business unit and is staffed by a team of 60 professionals, including six doctorates.
7. The CFA Institute surveyed a broad set of investors and analysts about XBRL and published a report. Several small fund managers and analysts have mentioned that XBRL would likely make them more competitive. For example, one respondent said, "Being able to get reliable data, on a timely basis, is critical to my ability to create and update valuation models" ([14], 4). Another respondent commented, "As a small investment manager, XBRL is a cost-effective way for me to build models so that I don't have to spend my time manually inputting documents" ([14], 17).
8. As an example, for a firm that adopts XBRL from the fiscal year ending December 31, 2009, its first XBRL-tagged 10-K report is filed in early 2010.
9. The Abel Noser database covers approximately 10 percent of all institutional trading activity ([36]).
10. This is because TAQ reports trades time-stamped to the nearest second. However, high-frequency traders execute trades in milliseconds. Consequently, hundreds—often, thousands—of trades appear to have the same time-stamp in TAQ. This renders Lee-Ready-type "tick" tests for inferring trade directions highly unreliable.
11. Note that we delete the institution-year-month with less than 100 orders and not the entire institution from our sample. Thus, if an institution has less than 100 orders in a few months, but more than 100 orders in other months, then only the months with less than 100 orders are eliminated. However, we obtain virtually identical results without imposing this restriction.

12. To operationalize this data screen, we only deleted the specific institution-stock pair if the number of trading days in a year is below the threshold instead of eliminating the entire institution. This is because an institution often administers multiple funds with different investment objectives (e.g., active versus passive buy-and-hold investment goals). Our approach is similar to studies in the mutual fund literature that focus only on actively managed funds from an institution's entire portfolio (e.g., [1]; [7]).
13. Our results continue to hold if we delete observations with earnings announcements in the (−5, +5) and (−10, +10) windows surrounding the 10-K announcement date.
14. We find similar results using the average daily dollar value of shares traded by the same institution-stock pair over days −90 to +1 as an alternative scaling variable.
15. We define yet another abnormal trading measure as the number of trades that take place in the event window divided by the sum of the number of trades in the pre-event window and the event window (labeled as ATRD). We also define a measure to capture the abnormal trading frequency during the event period (labeled as AFREQ). AFREQ is defined as the number of days during the three-day event window that an institution trades a firm's shares divided by the number of days the same institution trades the same firm's shares over both the pre-event and event windows. Our main results flow through using these alternative trading response measures, ATRD and AFREQ.
16. As mentioned earlier, a firm is included in our sample only if the firm is traded at least once by both a small institution and a large institution in each calendar year during our sample period. In a sensitivity check, we repeat our abnormal volume analysis by including only those observations for which there is at least one small institution and one large institution in the [−1, +1] window, and we obtain similar results.
17. We make the implicit assumption that information is priced within the 10-K event window the same way with or without XBRL adoption. It is, however, possible that XBRL reporting itself could change the nature of price formation within the 10-K window. To entertain this possibility, we include a term interacting ABS(10KCAR) with XBRL as an additional control variable. This way, the weight on ABS(10KCAR) could be different in the post-adoption time period from the pre-period. Since ABS(10KCAR) is the control for market reaction within the event window, this design provides the flexibility for the market reaction to be different in the two time periods. Our inferences are unchanged with this modification.
18. Appendix A provides additional details on how STYLE is calculated.
19. We run robustness checks assuming that the announcement period stretches up to four days, six days, eight days, and ten days after the 10-K filing date, and the results are similar.
20. Again, we repeat our analysis using a measure analogous to the TRADING variable, but based on share trading volume as opposed to dollar trading volume, and obtain nearly identical results.
21. The following example illustrates this point. Suppose the abnormal trading volume (AVOL) in the 10-K announcement window of an institution trading firm X's shares is ten in the pre-XBRL period, while AVOL in the 10-K release window of the same institution trading firm X's shares is 20 in the post-XBRL period. A trading volume-based measure will show an increase in abnormal volume, but our dichotomous measure (TRADING) will show no change (i.e., there is one observation where TRADING equals 1 in the pre-period and also one observation where TRADING equals 1 in the post-period). Furthermore, if the institution simply switches its trading coverage from firm X to firm Y after the XBRL adoption (i.e., AVOL is positive only for firm X in the pre-XBRL disclosure regime, while AVOL is positive only for firm Y in the post-XBRL regime), then, again, our TRADING metric will register no change from the pre- to the post-periods. In contrast, if AVOL is positive for firm X and zero or negative for firm Y in the pre-XBRL period, but AVOL is positive for both firms X and Y in the post-period, then the mean value of the TRADING metric will be higher in the post-

XBRL period. That is, there is one observation in the pre-period where TRADING equals 1, while there are two observations in the post-period where TRADING equals 1.

22. Probit estimation is often sensitive to the inclusion of a large number of fixed effects in the model ([25]). Therefore, we repeat this analysis using linear probability estimation (i.e., ordinary least squares [OLS]) and find qualitatively similar results.
23. Instead of clustering by firm and year, we cluster by firm and year-month because clustering by a variable with a small number of options can create inconsistency in standard error estimates ([34]; [24]). While our firm clusters have enough observations, our year clusters may have an insufficient number of options. Hence, we cluster by year-month to ensure that our standard error estimates are consistent and stable.
24. Another condition for causal identification in this DiD design is the assumption that small and large institutions' trading patterns would have continued on parallel trends and that they were not already converging in ways that would lead to the effects observed after the mandated adoption. Untabulated analysis shows that there is no difference in trading behaviors between small and large institutions prior to the XBRL adoption, suggesting that their trading patterns were not converging before the regulatory mandate.
25. The financial crisis of 2008 and 2009 likely affected large and small institutions differently. To ensure that our inferences are not sensitive to the crisis years, in untabulated analyses, we repeat our trading responsiveness tests involving AVOL, SPEED, and TRADING after eliminating 2008 and 2009 from our sample, and find similar results.
26. Additionally, we follow the established approach of identifying non-transient institutions using 13F data. Following the procedure outlined in [11], we classify institutions into three categories: transient, dedicated, and quasi-indexing. We label dedicated and quasi-indexing institutions as non-transient. Since 13F data report institutional holdings on a quarterly basis, we cannot implement our more powerful trading-based metrics to examine institutional activities on a day-by-day basis around 10-K filings. Instead, we rely on the change in institutional ownership before and after the 10-K release as a gauge of an institution's responsiveness to 10-K information. We analyze transient and non-transient institutions separately. We find that small transient institutions' responsiveness to 10-K reports (measured by the sensitivity of ownership change in response to the 10-K news) has increased to a greater extent compared to the change experienced by large transient institutions from the pre- to post-XBRL periods after controlling for various determinants of institutional ownership changes. In sharp contrast, we find no differential impact of XBRL on the sensitivity of ownership changes in response to 10-K news of small vis-à-vis large non-transient institutions.
27. For example, Baker et al (2010) infer mutual fund managers' stock picking skills from their ability to buy stocks that are about to enjoy high returns upon their upcoming earnings announcements and to sell stocks that are about to suffer low returns upon earnings announcements. Bhojraj et al (2012) use a similar measure to determine fund managers' stock picking skills.
28. The following example illustrates how STOCKPICKING 1 is computed. Suppose an institution-stock pair is involved in the following sequence of transactions during the three-day window centered on the firm's 10-K announcement date: (1) sells 300 shares at 1:00 p.m. on day −1 at a price of \$4.10, (2) buys 300 shares at 2:00 p.m. on day 0 at a price of \$4.30, and (3) buys 400 shares at 11:30 a.m. on day +1 at a price of \$4.50. Assuming the closing price of the stock on day +3 is \$5.00, STOCKPICKING 1 is calculated as follows: STOCKPICKING 1 $\equiv \{(0.2195 \times -1{,}230) + (0.1628 \times 1{,}290) + (0.1111 \times 1{,}800)\} \div 4{,}320 \equiv 0.0324$ or 3.24 percent.
29. Firm-level complexity is likely fairly sticky over time. Consequently, the tag counts of our sample firms are likely fairly stable over our sample period. Therefore, a measure based on tag counts in 2012 seems a reasonable proxy for complexity in both the pre- and post-adoption periods.

30. We run sensitivity tests using quartile instead of quintile as the break point and obtain similar results.
31. Since we are unable to uniquely identify institutions for the years 2011 and 2012, we use broker fixed effects for these two years instead of institution fixed effects.

REFERENCES

[1] Ali, A., X. Chen, T. Yao, and T. Yu. 2008. Do mutual funds profit from the accruals anomaly? Journal of Accounting Research. 46 (1): 1–26. https://doi.org/10.1111/j.1475-679X.2007.00263.x

[2] Anand, A., P. Irvine, A. Puckett, and K. Venkataraman. 2013. Institutional trading and stock resiliency: Evidence from the 2007–2009 financial crisis. Journal of Financial Economics. 108 (3): 773–797. https://doi.org/10.1016/j.jfineco.2013.01.007

[3] Angel, J. J., L. E. Harris, and C. S. Spatt. 2011. Equity trading in the 21st century. Quarterly Journal of Finance. 1 (1): 1–53. https://doi.org/10.1142/S2010139211000067

[4] Asthana, S., S. Balsam, and S. Sankaraguruswamy. 2004. Differential response of small versus large investors to 10-K filings on EDGAR. The Accounting Review. 79 (3): 571–589. https://doi.org/10.2308/accr.2004.79.3.571

[5] Baker, M., L. Litov, J. A. Wachter, and J. Wurgler. 2010. Can mutual fund managers pick stocks? Evidence from their trades prior to earnings announcements. Journal of Financial and Quantitative Analysis. 45 (5): 1111–1131. https://doi.org/10.1017/S0022109010000426

[6] Bhattacharya, U., C. Holden, and S. Jacobsen. 2012. Penny wise, dollar foolish: Buy-sell imbalances on and around round numbers. Management Science. 58 (2): 413–431. https://doi.org/10.1287/mnsc.1110.1364

[7] Bhojraj, S., Y. Cho, and N. Yehuda. 2012. Mutual fund family size and mutual fund performance: The role of regulatory changes. Journal of Accounting Research. 50 (3): 647–684. https://doi.org/10.1111/j.1475-679X.2011.00436.x

[8] Blankespoor, E., B. P. Miller, and H. D. White. 2014. Initial evidence on the market impact of the XBRL mandate. Review of Accounting Studies. 19 (4): 1468–1503. https://doi.org/10.1007/s11142-013-9273-4

[9] Bloomfield, R. 2002. The incomplete revelation hypothesis and financial reporting. Accounting Horizons. 16 (3): 233–243. https://doi.org/10.2308/acch.2002.16.3.233

[10] Bovee, M., A. Kogan, K. Nelson, R. P. Srivastava, and M. A. Vasarhelyi. 2005. Financial reporting and auditing agent with net knowledge (FRAANK) and extensible business reporting language (XBRL). Journal of Information Systems. 19 (1): 19–41. https://doi.org/10.2308/jis.2005.19.1.19

[11] Bushee, B. 2001. Do institutional investors prefer near-term earnings over long-run value? Contemporary Accounting Research. 18 (2): 207–246. https://doi.org/10.1506/J4GU-BHWH-8HME-LE0X

[12] Campbell, J. Y., T. Ramadorai, and A. Schwartz. 2009. Caught on tape: Institutional trading, stock returns, and earnings announcements. Journal of Financial Economics. 92 (1): 66–91. https://doi.org/10.1016/j.jfineco.2008.03.006

[13] Chambers, A. E., and S. H. Penman. 1984. Timeliness of reporting and the stock price reaction to earnings announcements. Journal of Accounting Research. 22 (1): 21–47. https://doi.org/10.2307/2490700

[14] Chartered Financial Analyst Institute Center for Financial Market Integrity (CFA Institute). 2007. XBRL Survey Report.. Charlottesville, VA: CFA.

[15] Chemmanur, T., S. He, and G. Hu. 2009. The role of institutional investors in seasoned equity offerings. Journal of Financial Economics. 94 (3): 384–411. https://doi.org/10.1016/j.jfineco.2008.12.011
[16] Chen, G., J. B. Kim, J. H. Lim, and J. Zhou. 2017. XBRL adoption and bank loan contracting: Early evidence. Journal of Information Systems. (forthcoming). https://doi.org/10.2308/isys-51688
[17] Clement, M. B. 1999. Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter? Journal of Accounting and Economics. 27 (3): 285–303. https://doi.org/10.1016/S0165-4101(99)00013-0
[18] Debreceny, R. S., S. M. Farewell, M. Piechocki, C. Felden, and A. Graning. 2010. Does it add up? Early evidence on the data quality of XBRL filings to the SEC. Journal of Accounting and Public Policy. 29 (3): 296–306. https://doi.org/10.1016/j.jaccpubpol.2010.04.001
[19] Debreceny, R. S., S. M. Farewell, M. Piechocki, C. Felden, A. Graning, and A. D'Eri. 2011. Flex or break? Extensions in XBRL disclosures to the SEC. Accounting Horizons. 25 (4): 631–657. https://doi.org/10.2308/acch-50068
[20] Debreceny, R. S., A. Chandra, J. J. Cheh, D. Guithues-Amrhein, N. J. Hannon, P. D. Hutchison, D. Janvrin, R. A. Jones, B. Lamberton, A. Lymer, M. Mascha, R. Nehmer, S. Roohani, R. P. Srivastava, S. Trabelsi, T. Tribunella, G. Trites, and M. A. Vasarhelyi. 2005. Financial reporting in XBRL on the SEC's EDGAR system: A critique and evaluation. Journal of Information Systems. 19 (2): 191–210. https://doi.org/10.2308/jis.2005.19.2.191
[21] Efendi, J., J. Park, and C. Subramaniam. 2016. Does the XBRL reporting format provide incremental information value? A study using XBRL disclosures during the voluntary filing program. Abacus. 52 (2): 259–285. https://doi.org/10.1111/abac.12079
[22] Goldstein, M., P. Irvine, E. Kandel, and Z. Wiener. 2009. Brokerage commissions and institutional trading pattern. Review of Financial Studies. 22 (12): 5175–5212. https://doi.org/10.1093/rfs/hhp083
[23] Gompers, P., and A. Metrick. 2001. Institutional investors and equity prices. Quarterly Journal of Economics. 116 (1): 229–259. https://doi.org/10.1162/003355301556392
[24] Gow, I. D., G. Ormazabal, and D. J. Taylor. 2010. Correcting for cross-sectional and time-series dependence in accounting research. The Accounting Review. 85 (2): 483–512. https://doi.org/10.2308/accr.2010.85.2.483
[25] Greene, W. 2004. The behavior of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects. Econometrics Journal. 7 (1): 98–119. https://doi.org/10.1111/j.1368-423X.2004.00123.x
[26] Harris, T. S., and S. Morsfield. 2012. An Evaluation of the Current State and Future of XBRL and Interactive Data for Investors and Analysts. White Paper Number Three, Columbia University Center for Excellence in Accounting and Security Analysis.
[27] Holden, C., and S. Jacobsen. 2014. Liquidity measurement problems in fast, competitive markets: Expensive and cheap solutions. Journal of Finance. 69 (4): 1747–1785. https://doi.org/10.1111/jofi.12127
[28] Hvidkjaer, S. 2008. Small trades and the cross-section of stock returns. Review of Financial Studies. 21 (3): 1123–1151. https://doi.org/10.1093/rfs/hhn049
[29] Ke, B., and S. Ramalingegowda. 2005. Do institutional investors exploit the post-earnings announcement drift? Journal of Accounting and Economics. 39 (1): 25–53. https://doi.org/10.1016/j.jacceco.2004.02.002
[30] Kim, J., B. Li, and Z. Liu. 2014. Who Benefits More from XBRL Adoption? Evidence from Breadth of Ownership. Working paper, City University of Hong Kong.
[31] Lee, C. M. C., and M. Ready. 1991. Inferring trade directions from intraday data. Journal of Finance. 46 (2): 733–746. https://doi.org/10.1111/j.1540-6261.1991.tb02683.x

[32] Miller, B. 2010. The effects of reporting complexity on small and large investors trading. The Accounting Review. 85 (6): 2107–2143. https://doi.org/10.2308/accr.00000001
[33] Nocera, J. 2004. Wall Street on the run. Fortune. (June 1). Available at: https://www1.warrington.ufl.edu/graduate/academics/msf/docs/speakers/prereading_McMahon_Ward0405FortuneMag.pdf
[34] Petersen, M. A. 2009. Estimating standard errors in financial panel data sets: Comparing approaches. Review of Financial Studies. 22 (1): 435–480. https://doi.org/10.1093/rfs/hhn053
[35] Plumlee, D. R., and M. A. Plumlee. 2008. Assurance on XBRL for financial reporting. Accounting Horizons. 22 (3): 353–368. https://doi.org/10.2308/acch.2008.22.3.353
[36] Puckett, A., and X. Yan. 2011. The interim trading skills of institutional investors. Journal of Finance. 66 (2): 601–633. https://doi.org/10.1111/j.1540-6261.2010.01643.x
[37] Securities and Exchange Commission (SEC). 2009. Interactive Data to Improve Financial Reporting. Available at: http://www.sec.gov/rules/final/2009/33-9002.pdf
[38] You, H., and X. Zhang. 2009. Financial reporting complexity and investor underreaction to 10-K information. Review of Accounting Studies. 14 (4): 559–586. https://doi.org/10.1007/s11142-008-9083-2

APPENDIX A
Variable Definition

Variable Definition

AVOL AVOL is computed as the average daily dollar value of shares of a firm traded by an institution over the three-day window centered on the firm's 10-K filing date (day 0) minus the average daily dollar value of shares of the same firm traded by the same institution over the pre-filing period of days -10 to -2. This measure is then scaled by the average daily dollar value of shares transacted by the same institution-stock pair over the days -10 to +1.

SPEED SPEED is computed as the total dollar volume of shares of a firm traded by an institution during the three-day period centered on the 10-K filing date (days _1 to +1), divided by the total dollar volume of shares of the same firm traded by the same institution over the seven-day period starting from the day before the filing date (i.e., days _1 to +5).

TRADING TRADING is an indicator variable that takes the value of 1 if AVOL is positive (that is, greater than zero); otherwise, it assumes the value of 0. This variable is defined only for institution-firm pairs that existed at least once in both the pre-XBRL period and the post-XBRL period.

STOCKPICKING 1 STOCKPICKING 1 is a weighted average return measure using all transactions related to an institution-firm pair over the three-day event window centered on the firms' 10-K filing date (days -1 to +1), where the dollar volume of trades associated with each transaction acts as the weight. For each transaction in the 10-K event window, a weighted return measure is first computed by multiplying the return from that transaction with the corresponding dollar value of shares traded in the transaction and an indicator of +1 for purchase or -1 for sale, where return is based on the actual transaction price and the closing price on day +3. The weighted return measures of all the transactions in the event window are then aggregated and scaled by the cumulative dollar volume of all trades in the window to arrive at a weighted average return measure for each institution-firm pair.

STOCKPICKING 2 STOCKPICKING 2 is calculated the same way as STOCKPICKING 1, except that the return is based on the actual transaction price and closing price on day +1.

XBRL XBRL is an indicator variable. It takes the value of 1 if a company whose stocks are traded by our sample institutions files its 10-K in a given year using XBRL-tagged data, and it is coded 0 if the company files using the HTML format. XBRL is defined at the stock-year level.

SMALL SMALL is an indicator variable. SMALL takes the value of 1 if a sample institution in a given year is classified as a small institution, and it is coded 0 otherwise. SMALL is defined at the institution-year level.

LOG(MV) Natural logarithm of market capitalization of the firm, defined as the number of shares outstanding multiplied by the closing price, both measured as of the end of the fiscal year of the 10-K announcement.

MTB Market-to-book ratio of the firm, measured as the market value of common equity divided by the book value of common equity, both measured as of the end of the fiscal year of the 10-K announcement.

MOM The stock momentum (MOM) of a company, measured by daily stock returns compounded over a 90-day period ending one day before the firm's 10-K release date minus daily market returns compounded over the same period.

WC Natural logarithm of the number of words in a firm's 10-K report.

ABS(10KCAR) ABS(10KCAR) is measured as a decile ranking of the absolute value of cumulative market-adjusted returns over a three-day period centered on the firm's 10-K filing date.

ABS(EACAR) ABS(EACAR) is measured as a decile ranking of the absolute value of cumulative market-adjusted returns over a three-day window centered on the firm's annual earnings announcement date.

AFTEAD AFTEAD is the number of calendar days between the earnings announcement and the 10-K announcement.

ONTIME ONTIME is an indicator variable that takes the value of 1 if the 10-K is filed within one day from the expected filing date, and it is 0 otherwise. The expected filing date is the same day of the month of last year's 10-K filing.

STYLE The trading style of an institution captures its propensity to trade in the direction of the daily prevailing market return of a particular stock. We follow Anand et al. (2013) and classify a buy-order as $Volume_{With}$ if the stock return for the day is positive, and as $Volume_{Against}$ if the stock return for the day is negative. The converse applies for sell-orders. For each institution, STYLE is calculated as follows based on the aggregate dollar trading volume with and against the contemporaneous daily stock returns in the year prior to the 10-K filing date:

$$Trading\ Style = \frac{\sum Volume_{With} - Volume_{Against}}{\sum Volume_{With} + Volume_{Against}}.$$

COST The trading cost of an institution, calculated as the sum total of the trading commission, fees, and taxes paid in a year, divided by the number of shares traded in that year.

COMPLEX COMPLEX takes the value of 1 if the number of tags in a firm's 2012 10-K filing is above the top quintile of our sample, and it takes the value of 0 if the tag count falls below the top quintile. COMPLEX is defined at the firm level.