Digital Insiders and Informed Trading Before Earnings Announcements

Henk Berkman [a]
Jonathan Jona [b*]
Gladys Lee [c]
Naomi Soderstrom [d]

Draft date: February 1 2021

[a] The University of Auckland Business School & The University of Sydney, Owen G Glenn Building, Auckland 1010, New Zealand. Email: h.berkman@auckland.ac.nz
[b] Freeman School of Business, Tulane University. 7 McAlister Dr, S, New Orleans, LA 70118. Email: jjona@tulane.edu
[c] The University of Melbourne, 198 Berkeley Street, Parkville, 3010, Victoria Melbourne. Email: gladys.lee@unimelb.edu.au
[d] The University of Melbourne, 198 Berkeley Street, Parkville, 3010, Victoria Melbourne. Email: naomiss@unimelb.edu.au

[*] Corresponding author.
We are grateful to Jackie Cook from CookESG Research for her help in extracting the cybersecurity excerpts. We gained valuable insights from Angelo Angelis, Mary Barth, Kasper Jønsson, Katherine Schipper, Devin Shanthikumar (discussant), workshop participants at Aarhus University, Copenhagen Business School, LaTrobe University, Lingnan University, London School of Economics, Tel-Aviv University, University of Canterbury, University of Otago and University of Technology Sydney, and conference participants at the FARS conference (2020). We thank Marco Eugster, James Kavourakis, and Rachel Solano for research assistance.

Funding: Jona acknowledges the financial support provided by the University of Melbourne through an Early Career Researcher grant.

Digital Insiders and Informed Trading Before Earnings Announcements

Abstract

In addition to growing risk from hackers stealing customer information, an increasingly common cybersecurity risk for firms stems from digital insiders – hackers who target corporations to obtain non-public corporate information for illegal trading. We propose and validate two firm-specific measures of cybersecurity risk mitigation based on textual analysis of 10-Ks to proxy for the organization's ability to reduce the probability of digital insider trading. We find that a larger share of new earnings information is incorporated into prices prior to earnings announcements for firms with low cybersecurity risk mitigation scores and that pre-announcement trading by short sellers is more predictive of earnings surprises for these firms. Also suggestive of informed trading, for firms with lower cybersecurity risk mitigation scores we find an increase in stock and option trading volume and higher intraday price volatility during several weeks prior to earnings announcements.

JEL classification: G14, G18, K24, M48, M41

Keywords: cybersecurity risk mitigation, probability of informed trading, liquidity, cybersecurity, cybersecurity risk disclosure, digital insiders, bid ask spread, private information, hacking, price jump ratio, textual analysis, short selling.

Digital Insiders and Informed Trading before Earnings Announcements

1. Introduction

Between 2013 and 2014, a group of hackers, "FIN4", illegally obtained data for trading purposes from more than 100 U.S. companies, systematically targeting employees who might possess value-relevant non-public information, such as C-level executives, legal counsel, and risk and compliance personnel.[1] According to Joseph Carson, Chief Security Scientist and Advisory CISO at Thycotic, rather than installing malware or ransomware to obtain a payout, such 'digital insiders' use methods such as phishing emails or trojan horses to obtain access to privileged accounts and information (Carson 2017).

In a 2019 blog post, Carson comments that, "For the cybercriminal, the goal is NOT to install malicious malware or disruptive ransomware forcing the company to pay-out, in fact these cyber criminals do not even steal the data or threaten to disclose it. In common with nation state actors, cyber criminals do not want to be detected, and so employ the same techniques – their goal is financial gain, and to do this they need to remain hidden from their unsuspecting victims."[2] Once the cybercriminals insinuate themselves into the firm's system, they install surveillance tools to gather valuable undisclosed information, which can subsequently be exploited in the stock market.

[1] https://www.computerworld.com/article/2853697/fireeye-suspects-fin4-hackers-are-americans-after-insider-info-to-game-stock-market.html and http://securityaffairs.co/wordpress/38118/cyber-crime/sec-investigates-fin4-hackers.html. Cyber criminals have also successfully targeted media firms (https://www.sec.gov/news/pressrelease/2015-163.html), law firms, (https://www.welivesecurity.com/2017/05/11/hackers-stole-information-law-firms-made-millions-insider-trading-fined-9-million/) and advisory firms (https://www.theguardian.com/business/2017/sep/25/deloitte-hit-by-cyber-attack-revealing-clients-secret-emails) to steal non-public information about mergers and acquisitions or earnings announcements. We note that trades based upon hacked information can be made by the hackers themselves or by others who receive the information derived from hacking.

[2] See https://www.informationsecuritybuzz.com/expert-comments/australian-hacker-jailed-for-insider-trading/.

In 2020, top U.S. federal agencies confirmed that they were victims of a world-wide hacking campaign perpetrated through illegitimate software updates for SolarWinds, a tool that is widely used in both government and corporate settings for network monitoring. Investigation of the incident revealed that hackers deployed the compromised software approximately 9 months prior to its detection.[3] In addition to hacking company systems from the outside, digital insiders can be current or former company insiders (such as employees of the target firm)[4] or can gain access to sensitive information from such company insiders.[5]

While these activities are potentially very damaging to the integrity of financial markets, due to their secretive nature it is difficult to establish their total impact. This opacity also makes it difficult to evaluate the effectiveness of corporate strategies to reduce the potential for digital insiders to gain access. In this paper, we examine whether firm cybersecurity risk mitigation affects the information content of prices, which could be influenced by digital insider activities. Focusing on earnings announcements as significant information events, we explore the relation between firm cybersecurity risk mitigation and the extent to which private information is traded on and is reflected in prices prior to the announcements.

---

[3] https://www.wsj.com/articles/solarwinds-discloses-earlier-evidence-of-hack-11610473937?mod=article_inline . Some digital insiders can remain undetected for many years. IT consultant Steven Oakes was sentenced to three years imprisonment related to insider trading based upon information he gained through hacking of network login credentials for the financial publisher Port Phillip Publishing. On 70 occasions between January 2012 and February 2016, Oakes made trades to profit off of unpublished buy recommendations for shares on the Australian Stock Exchange (https://www.zdnet.com/article/aussie-hacker-jailed-for-unauthorised-access-and-insider-trading/#ftag=RSSbaffb68).

[4] In 2016, a former Expedia IT specialist remotely hacked into computers and email accounts of senior Expedia executives and made highly profitable trades in Expedia securities ahead of company announcements (https://cdn.arstechnica.net/wp-content/uploads/2016/12/ly-complaint-sec.pdf).

[5] A report by RedOwl (2017) states that "Sophisticated threat actors use the dark web to find and engage insiders to help place malware behind an organization's perimeter security". See https://www.nationalinsiderthreatsig.org/itrmresources/RedOwl%20Report-Monetizing%20The%20Insider%20Through%20The%20Dark%20Web.pdf. We do not differentiate between informed trading by digital insiders who are company insiders versus those who are external to the firm.

We develop two cybersecurity risk mitigation measures based on textual analysis of 10-K disclosures starting in 2012. This timeframe is driven by SEC guidance in 2011, which requires companies to include material information related to cybersecurity risk in their periodic filings (*CF Disclosure Guidance: Topic No. 2 Cybersecurity*). We argue that 10-K disclosures contain information about how firms address and mitigate cybersecurity risks and are unlikely to provide information that exposes them to additional cyber-related risk from hackers.[6] Our first measure is a simple count of the number of words in cybersecurity-related excerpts from the 10-K reports. We identify cybersecurity-related excerpts using cybersecurity-related words based on a dictionary of cybersecurity terminology from the glossary of the National Initiative for Cybersecurity Careers and Studies (NICCS) and a report on laws relating to cybersecurity prepared by the Congressional Research Service (Fischer 2014). Our second measure is based upon a narrower definition of cybersecurity risk mitigation, representing firm strategies, policies, and mechanisms that address cybersecurity risks. We develop a mitigation dictionary (See Appendix A.1) based upon textual analysis of these excerpts. Our second measure is a count of these mitigation words or word combinations within the cybersecurity-related excerpts.

We validate both cybersecurity risk mitigation measures by showing that firms disclosing more cybersecurity words or more cybersecurity risk mitigation words have a lower likelihood of future cybersecurity data breaches. We also find that there is an increase in the use of cybersecurity words and cybersecurity risk mitigation words for firms in the two years following data breaches at peer firms. Our measures perform better than four alternative measures of the quality of firm-specific cybersecurity policies, including tone of the cybersecurity disclosures (Loughran and

[6] SEC (2011) states that, "federal securities laws do not require disclosure that itself would compromise a registrant's cybersecurity." This interpretation is consistent with the result in Berkman et al. (2018), who find positive market valuations for a cybersecurity awareness measure based upon 10-K disclosures.

McDonald 2011),[7] information technology (IT) capital expenditures (Ashraf and Sunder 2018; Ashraf et al. 2020), presence of an IT or risk board committee (Higgs et al. 2016), and the number of IT executives in top management of the firm (Kwon et al. 2013).

After validating our measures, we test the prediction that firms with higher cybersecurity risk mitigation scores are less likely to experience leakage of inside information due to the activities of cyber criminals. Our empirical tests examine the impact of cybersecurity risk mitigation on the probability that private information will be traded on and be reflected in prices before earnings announcements. We focus on earnings announcements because they are frequent information events where value-relevant information is made public. In addition, much of the documented activity of digital insiders concerns trading preceding earnings announcements.

Our first test employs the price jump ratio (Weller 2018) to investigate the relative pre-announcement information content of prices. The price jump ratio (PJR) is defined as the cumulative abnormal return (CAR) in a short window around the earnings announcement relative to the CAR for the same stock over a longer pre-announcement window ending on the same day as the short earnings announcement window. Consistent with our expectation, we find that firms with lower cybersecurity risk mitigation scores have a lower price jump ratio, indicating that a relatively larger proportion of earnings information is incorporated into prices prior to the earnings announcement.

Our second test examines the impact of a firm's cybersecurity risk mitigation on the relative amount of short selling before earnings announcements due to informed trading. This test is motivated by several charges brought by the SEC, where hackers short-sold stocks just before

[7] We thank one of the reviewers for this suggestion.

firm disclosure of disappointing earnings news.[8] Consistent with our hypothesis, we find that the ability of informed traders (proxied by the relative trading activity of short sellers) to predict earnings surprises is greater for firms with lower levels of cybersecurity risk mitigation. Further analysis suggests that the window over which digital insiders trade stretches over several weeks prior to the earnings announcement.

Our final set of tests uses measures of information asymmetry to capture the impact of cybersecurity risk mitigation on the probability of informed trading in the pre-earnings announcement period (e.g., Ahern 2020; Akey et al. 2020; Collin-Dufresne and Fos 2015; Kacperczyk and Pagnotta 2019). Consistent with our hypothesis, we provide evidence that in the weeks leading up to earnings announcements, firms with lower cybersecurity risk mitigation scores experience larger abnormal stock trading volume, larger abnormal option trading volume, and higher intraday price volatility. We also provide consistent evidence for alternative measures of the probability of informed trading.

Our study makes several contributions to the literature. We contribute to the literatures on cybersecurity risk disclosures and informed trading by providing evidence that cybersecurity disclosures in 10-K filings explain variation in the probability of private information leakage. Specifically, we show that cybersecurity risk mitigation scores based on a firm's 10-K disclosures are informative and help to explain pre-earnings announcement price formation. This extends the literature on the informativeness of 10-K disclosures and in particular, risk disclosures (e.g., Campbell et al. 2014; Hope et al. 2016). Our evidence indicates that the disclosures are not boilerplate, but that they contain value-relevant information. With increased recognition of risks associated with hackers infiltrating firms and obtaining private information for illegal trading, our

[8] For example, https://www.nbcnews.com/business/business-news/u-s-charges-9-insider-trading-based-hacked-press-releases-n407771.

findings become more relevant. Regulators face calls for improved cybersecurity disclosure requirements from within the SEC.[9] Auditors face increased demands to understand potential weaknesses in the systems that support the financial reporting process (Hamm 2019). Our results suggesting the effectiveness of cybersecurity mitigation in curtailing digital insider trading provide insights to these stakeholders and other market participants[10] as well as to the academic literature (Gordon et al. 2015; Ferraro 2014; Selznick and LaMacchia 2016).

We also contribute by introducing and validating two simple firm-specific measures of the quality of cybersecurity mitigation policies. Prior studies have developed measures of cyber risk based upon disclosures that capture the overall salience and awareness of cybersecurity risk (see Berkman et al. 2018), or specific cybersecurity policies (see Ashraf et al. 2020; Kwon et al. 2013; Higgs et al. 2016). Our first measure, based on the total number of words in cyber excerpts, addresses cybersecurity mitigation more generally. Our second measure, based on our mitigation-word count, is specifically focussed on mitigation words within the excerpts. Both measures are constructed based on publicly available information in 10-Ks and can be applied in future research to capture a firm's cybersecurity risk environment or can be used as a proxy of informed trading stemming from cybercrimes.

Finally, our paper contributes to the literature on the relation between earnings and returns by providing evidence that informed digital insiders play a role in the way that earnings information gets impounded into prices in the period before it is released (e.g., Ball and Shivakumar 2008; Bushee et al. 2010; Huang and Skantz 2016). Ball and Shivakumar (2008) find

[9] See the statements from Commissioner Kara M. Stein (https://www.sec.gov/news/public-statement/statement-stein-2018-02-21) and Commissioner Robert J. Jackson Jr (https://www.sec.gov/news/public-statement/statement-jackson-2018-02-21), who argue that existing cybersecurity disclosure requirements are not sufficient.

[10] https://www.csoonline.com/article/3260006/data-breach/secs-new-cybersecurity-guidance-falls-short.html, https://www.nytimes.com/2018/03/05/business/dealbook/sec-cybersecurity-guidance.html.

that at the time of earnings announcement, most information is already impounded into prices. We contribute to this literature on information asymmetry in the pre-earnings announcement period and price formation by showing that price informativeness in the pre-announcement period differs among firms based upon their cybersecurity risk management.

2. Literature review and hypothesis development

2.1 Prior literature on cybersecurity

Research on cybersecurity in accounting and finance can be split into two streams. The first stream provides evidence on the impact of cybersecurity disclosures on market valuations. In the period before the SEC guidance on disclosure of cybersecurity risk (SEC 2011), Gordon et al. (2010) find higher market valuations for the small proportion of firms that voluntarily disclosed cybersecurity risks. Subsequent to the SEC guidance and using a broad sample of firms, Berkman et al. (2018) find that more informative cyber disclosures are associated with higher market valuations.

The second stream of research examines the consequences of cybersecurity events. Several articles find evidence that positive cybersecurity events such as IT security investments (Bose and Leung 2013; Chai et al. 2011; Im et al. 2001) and/or creation of a Chief Information Officer position (Chatterjee et al. 2001) are associated with higher stock prices. Relatedly, Kwon et al. (2013) find that greater total IT executive compensation is associated with a lower likelihood of information security breaches. Firms also suffer fewer security breaches when they have stronger internal controls (Westland 2018) and when there is a higher quality relationship between the internal audit and the information security function (Steinbart et al. 2018).

Studies that examine the consequences of negative cybersecurity events generally find evidence of negative market reactions to the events.[11] For example, research indicates that announcements of software vulnerability (Telang and Wattal 2007), IT products containing viruses (Hovav and D'arcy 2005), and cybersecurity breaches are associated with negative market reactions (Amir et al. 2018; Cavusoglu et al. 2004; Gordon et al. 2011; Kamiya et al. 2018; Modi et al. 2015; Yayla and Hu 2011), although Hilary et al. (2016) find a smaller reaction for breaches than for asset impairments. Bianchi and Tosun (2019) find that firms experience a decrease in liquidity and that daily excess returns are lower following the revelation of a first-time corporate hacking event. Huang and Wang (2020) find evidence that firms with reported data breaches have higher loan spreads and that their loans require more collateral. There is also evidence that some stock market participants obtain early notice of impending breach announcements and profit from short-selling (Mitts and Talley 2019) or insider trading (Lin et al. 2020). Akey et al. (2020) examine trading by hackers who illegally accessed earnings information by hacking major newswire services. They find that prior to its public release, private information traded upon by digital insiders is impounded in stock prices. These authors also show that the trading activity by hackers sharply increases order flow and bid ask spreads, which is consistent with traditional models of market microstructure.

While the above studies highlight the need for firms to actively mitigate cybersecurity risks, there is comparatively little research on the impact of firm-specific cybersecurity risk mitigation. One exception is a study by Wang et al. (2013) who find that firms disclosing security risk factors with risk-mitigation themes are less likely to have future breach announcements. We

[11] Spanos and Angelis (2016) provide a comprehensive review of the stock market impacts of security events.

extend this literature by providing evidence that firm cybersecurity risk mitigation strategies reduce the probability of informed trading by digital insiders prior to earnings announcements.

2.2 Hypotheses development

2.2.1 Cybersecurity mitigation and price informativeness

A key aspect of our study is investigating whether cybersecurity risk mitigation impacts how earnings-related information is impounded into stock prices. Early research on price formation employs a measure of intra-period timeliness (IPT) to investigate the speed with which information is impounded into prices during an earnings quarter (Alford et al. 1993; Ball and Brown 1968; Butler et al. 2007; McNichols 1984). Fast price discovery implies that the end-of-quarter perfect foresight price level is attained early in the quarter.

Weller (2018) extends this literature by introducing the price jump ratio (PJR). The PJR measures the share of new earnings information incorporated into the price before the earnings announcement of a specific stock. The PJR is defined as the price change in a short window around the earnings announcement relative to the price change for the same stock over a longer pre-announcement window ending on the same day as the short earnings announcement window. A high PJR, indicating a large announcement price change relative to the pre-announcement price change, is consistent with little pre-announcement informed trading. In contrast, a low PJR, indicating a small announcement price change relative to the pre-announcement price change, suggests aggressive informed trading in the pre-announcement period.

Based on the conjecture that firms providing more extensive cybersecurity risk-mitigation disclosure are more likely to have taken measures to manage that risk, we hypothesize that informed trading by hackers is less likely in firms with high cybersecurity risk mitigation scores.

For these firms, we expect a higher PJR reflecting a larger proportion new information being incorporated into price following the earnings announcement.

H1: Ceteris paribus, earnings announcement price jump ratios are positively associated with cybersecurity risk mitigation.

2.2.2 Cybersecurity mitigation and likelihood of informed short selling

Digital insiders gain access to information that may relate to both positive and negative earnings surprises. Informed traders can benefit from this information through long or short selling in advance of the earnings announcement. We focus on short selling because, in contrast to trading on the long side, there is daily data on trading by well-informed investors. In a wide variety of settings, research provides evidence consistent with short sellers anticipating future information releases by exploiting private information. For example, there is greater short selling activity in the days leading up to downgrades by analysts (Christophe et al. 2010) and insider sales (Chakrabarty and Shkilko 2013; Khan and Lu 2013). Karpoff and Lou (2010) find that abnormal short selling increases in the period before disclosure of misrepresentations. In a cybersecurity-related setting, Mitts and Talley (2019) find evidence that prior to firm breach announcements, informed traders take short positions against the hacked firms. Studies also show that short interest increases prior to the announcement of private placements in which hedge funds are involved (Berkman et al. 2016) and that short sellers are able to profitably exploit material non-public information arising from the syndicated lending process (Massoud et al. 2011).

Weak cybersecurity risk mitigation by firms should increase the probability that hackers can obtain (and exploit or sell) private information related to earnings. As a result, with a higher probability of information leakage, there should be increased pre-announcement trading by short sellers in firms with lower cybersecurity risk mitigation when earnings surprises are negative.

H2: Ceteris paribus, short selling in the period before earnings announcements is more predictive of earnings surprises for firms with low scores on cybersecurity risk mitigation.

2.2.3 Cybersecurity mitigation and the probability of informed trading

Theoretical models of the cost of liquidity typically assume that one set of traders provides liquidity via quotes or limit orders and another set of traders initiates trades for liquidity or for informational reasons (Holden et al. 2014; Huang and Stoll 1996). These models typically posit that the spread provides suppliers of liquidity with compensation for: 1) adverse selection costs (Easley and O'Hara 1987; Glosten and Milgrom 1985; Kyle 1985); 2) order processing costs (Roll 1984); and 3) inventory holding costs (Amihud and Mendelson 1980; Ho and Stoll 1981; Ho and Stoll 1983). Building on these models, several papers attempt to measure components of the bid-ask spread, in particular the component related to adverse selection risk (e.g., Glosten and Harris 1988; Hasbrouck 1991; Lin et al. 1995; Stoll 1989).

However, recent studies find that traditional measures of adverse selection based on spreads fail to capture the presence of informed trading when informed investors strategically time their trading to occur when stocks are most liquid and because of order splitting to minimize price impact (e.g., Ahern 2020; Collin-Dufresne and Fos 2015; Kacperczyk and Pagnotta 2019). For example, similar to Ahern (2020), Kacperczyk and Pagnotta (2019) examine whether measures of information asymmetry reveal the trading of privately informed trading based on a sample of trades documented in the SEC's insider trading investigations. These authors find that standard measures of adverse selection generally perform poorly in their sample, although they provide evidence that informed trading activities are revealed by increased volume in the stock and options market and by increased intraday volatility in the stock market. In a study closely related to ours, Akey et al. (2020) examine trading by hackers of major newswire services. They find that volume on the stock

market and options market were higher when hackers were active. Moreover, consistent with the traditional microstructure literature, they find that effective spreads and realized spreads increase in reaction to this informed trading.

Firms with low cybersecurity mitigation are more susceptible to being hacked or infiltrated by digital insiders, with a concomitant higher likelihood of digital insiders trading on information about forthcoming earnings news. Building on the prior research findings that adverse selection is positively correlated with volume and volatility, we conjecture that in the period before earnings releases, firms with low cybersecurity risk mitigation are associated with higher trading volume and volatility.

H3: Ceteris paribus, stock volume, option volume, and intraday volatility in the pre-earnings announcement period are negatively associated with cybersecurity risk mitigation.

We note that increased trading volume prior to earnings announcements could be associated with uninformed traders taking cues from increased trading by hackers and "leaning with the wind" (Van Kervel and Menkveld 2019). In this case, while the hackers may not have direct responsibility for all of the market impact of their trading, it is their trading behavior that results in accelerated revelation of the earnings information in the pre-announcement period.[12]

[12] There is tension in our hypotheses stemming from the potential of weak cybersecurity risk mitigation to reduce incentives for market participants to perform in-depth analysis of firm performance prior to earnings announcements. Consistent with the argumentation in Weller (2018), if a firm has weak cybersecurity policies, it is possible that a hacker has already made trades such that fundamental information has been incorporated in price. As a result, the expected benefits for analysts performing fundamental analysis and their incentives to engage in this analysis are reduced for firms with weak cybersecurity policies, increasing the PJR and lowering information asymmetry.

3. Sample, variables and descriptive statistics

3.1 Sample

Our sample period starts with fiscal year 2012, the first fiscal year subsequent to the SEC guidance on cybersecurity risk disclosure. For fiscal years 2012–2018, we construct two cybersecurity risk mitigation measures (described below) using cybersecurity-related excerpts from 10-Ks of Russell 3000 firms. After merging the cybersecurity risk mitigation data with Compustat and CRSP, our sample includes 18,529 firm-year observations, with cybersecurity risk mitigation scores for 3,209 firms.

To identify 'event day 0', the first day after the quarterly earnings announcement that the closing price reflects the new earnings information, we use the earnings announcement date and time from the I/B/E/S database. We adjust event day 0 for after-hours earnings announcements. In addition, after merging I/B/E/S and Compustat databases using the linking table in WRDS and requiring the best match score, we require that the earnings announcement date be the same in both databases. As a result of these requirements, our main sample of earnings announcements contains 60,862 unique events, for 3,010 firms. We merge the short sales data and trade and quote data in DTAQ with our sample of earnings announcements using historical tickers based on the linking table in WRDS. For these tests our sample is reduced to 49,997 earnings announcements for 2,432 unique firms.

3.2 Variables

3.2.1 Cybersecurity risk mitigation measures

We construct our cybersecurity risk mitigation measures using excerpts from 10-K disclosures containing words or phrases that directly relate to cybersecurity themes. Cyber

disclosures are provided throughout the 10-K and are not restricted to Item 1A (Berkman et al. 2018; Gordon et al. 2015; SEC 2011, 2018). The primary areas of the 10-K include management's discussion and analysis of financial condition and results of operation (MD&A), description of business, description of legal proceedings, and Item 1A, Risk Factors.To identify cyber-related excerpts, we developed a keyword list using a glossary of common cybersecurity terminology from the National Initiative for Cybersecurity Careers and Studies (NICCS) and a report on laws relating to cybersecurity prepared by the Congressional Research Service (Fischer 2014). We incorporated the cyber-related keywords and phrases into a computer-based disclosure mapping logic to develop an initial corpus of cybersecurity disclosures. We refined this dictionary through an iterative process of testing the original list against samples of disclosures from a variety of industry groupings. In this process, the focus was to prune false positives while minimizing the risk of false negatives. Based upon this process, we have a corpus of excerpts from which we develop measures reflecting the firm's cybersecurity risk mitigation activities. Our first measure of cybersecurity mitigation is the total word count of these cybersecurity excerpts (*NCyberWords*).

To create a score that more specifically reflects cybersecurity risk mitigation, we identify a list of mitigation-related words (see Appendix A.1). This list was developed by a team consisting of one of the authors on this project and two independent research assistants. The team members independently identified mitigation-related words and phrases in the cyber excerpts. The team then resolved any discrepancies and refined the dictionary. Although many of these words (e.g., insurance and training) are not uniquely related to cybersecurity and thus could be generalized to other themes, we examine usage of these words solely within the context of cybersecurity (i.e., in the cybersecurity-related excerpts). This provides confidence that the disclosed mitigation measures relate specifically to cybersecurity and not to a more general risk management strategy

by the firm. For the measure *NMitigation*, each mitigation word / word combination within an excerpt receives a score of one and scores are tallied across all excerpts. Appendix A.2 provides examples of 10-K excerpts containing cybersecurity risk mitigation themes. Section 4.1 provides validation tests of these measures.

3.2.2 Price Jump Ratio

For our test of Hypothesis 1, the measure of PJR for the earnings announcement of stock *i* in quarter *t* is defined as:

$$PJR_{it} = CAR_{it}(0,2) / CAR_{it}(-21,2) \quad (1)$$

Similar to Weller (2018), we estimate cumulative abnormal returns (*CAR)* relative to a Fama and French (1992) three-factor model, using daily returns over a 365-calendar day window ending 90 days before the earnings announcement (we require at least 63 valid preceding trading days). We use this benchmark model to estimate the CAR over a total announcement period that starts on trading day -21 (about 1 month) and ends 2 trading days after the earnings announcement to ensure that prices fully reflect the new information. We similarly estimate the CAR over the announcement return window, which starts on day 0 and ends 2 days after the earnings announcement.[13]

To address the problem of a near-zero denominator in the PJR, we follow Weller (2018) and only retain observations where the absolute value of *CAR*(-21,+2) is larger than the daily return volatility in the preceding month multiplied with the square root of 24. By excluding observations with small denominators (i.e., *CAR*(-21,+2) is close to 0), we exclude observations with low signal-

[13] In robustness tests, we present results for a 10-day pre-announcement window. Note that our earnings announcement window starts on day 0 rather than day -1 as in Weller (2018). We adjust earnings announcement dates for after-hours announcements, whereas Weller (2018) uses Compustat earnings announcement dates, which are not adjusted for after-hours announcements.

to-noise ratios that are also non-events from the perspective of informed traders (see Weller 2018).[14] In robustness tests, we present results that use different exclusion cut-offs for *CAR*(-21,+2) as well as a trading volume-based measure of PJR.

3.2.3 Abnormal short-selling measure

For our test of Hypothesis 2, following prior literature (Christophe et al. 2004; Engelberg et al. 2012), we measure the level of daily short selling for firm *i* on day *t* as the daily number of shares sold short as a proportion (in percent) of total volume traded:

$$SHVOL_{i,t} = \frac{\text{Number of shares sold short}_{i,t}}{\text{Trading volume}_{i,t}} \times 100 \qquad (2)$$

For each earnings announcement, we calculate abnormal short selling for day -20 through day -1, using the mean and standard deviation based on the 20 trading days from day -50 to day -31. Abnormal short selling on day *t* relative to the earnings announcement day for quarter *q* for stock *i* is defined as follows:

$$ASHVOL_{i,q,t} = \left(SHVOL_{i,q,t} - M\_SHVOL_{i,q}\right) / S\_SHVOL_{i,q} \qquad (3)$$

Where:

$ASHVOL_{i,q,t}$	=	abnormal short selling on day *t* relative to quarter *q*'s earnings announcement day (day 0) for stock *i*.
$SHVOL_{i,q,t}$	=	actual short selling activity on day *t* relative to quarter *q*'s earnings announcement day for stock *i*.
$M\_SHVOL_{i,q}$	=	the mean daily short selling activity, measured over the 20-day period from day -50 to day -31 relative to quarter *q*'s earnings announcement day for stock *i*.
$S\_SHVOL_{i,q}$	=	the standard deviation of daily short selling activity, measured over the 20-day period from day -50 to day -31 relative to quarter *q*'s earnings announcement day for stock *i*.

[14] If stock prices move only in response to announcement news, the price jump ratio should be bounded in the [0, 1] interval. However, the empirical price jump ratio measure can be outside this interval because of idiosyncratic price movements unrelated to announcement news adding noise to the numerator and denominator. Our design choices are based on Weller (2018) and are consistent with recent papers using the PJR such as Chen et al. (2020) and Cao et al. (2020).

In our tests of hypothesis 2, we initially focus on the average daily abnormal short selling activity in the month before every earnings announcement, $ASHVOL(-20,-1)_{i,t}$.

3.2.4 Pre-announcement probability of informed trading

Our tests of Hypothesis 3 relate cybersecurity risk mitigation to the probability of informed trading in the pre-earnings announcement period and we initially focus on stock volume, option volume, price range and intraday volatility. Stock trading volume ($StkVol_{i,t}$) is measured as the natural log of the number of traded shares for firm *i* on day *t*. Option trading volume ($OptVol_{i,t}$) is measured as the natural log of the number of traded put and call options for firm *i* on day *t*. Price range ($Range_{i,t}$) is measured as the difference between the highest trade price minus the lowest trading price for firm *i* on day *t*. Intraday volatility (*Volatility*) is the intraday trade-based volatility during trading hours for firm *i* on day *t*.

We follow the same process as before and use the mean and standard deviation of the respective variables over the 20 trading days from day -50 to day -31 to standardize the daily measures in the month before the earnings announcement. We then calculate the average daily abnormal values for these variables for the 20 days before every earnings announcement to obtain $AStkVol_{i,t-20,t-1}$, $AOptVol_{i,t-20,t-1}$, $ARange_{i,t-20,t-1}$, and $AVolatility_{i,t-20,t-1}$.

3.3 Descriptive statistics

Table 1 Panel A presents descriptive statistics for the main variables in our study. The raw wordcount of cybersecurity-related excerpts ranges from 0 to 15,994, and the mean (median) wordcount of cybersecurity-related excerpts is 1,032 (712). The raw number of cybersecurity risk mitigation words in the cybersecurity excerpts ranges from 0 to 483, with the mean (median) number of cybersecurity risk mitigation words in the cybersecurity-related excerpts being 18 (11).

To reduce the impact of the right skewness of these variables, our empirical tests use the natural log of one plus the wordcount of cybersecurity-related excerpts (*CyberWords*), and the natural log of one plus the number of cybersecurity risk mitigation words (*Mitigation*).

We winsorize the dependent variables used in our tests of the impact of mitigation measures on pre-earnings announcement information leakage (i.e., *PJR*, $ASHVOL(-20,-1)_{i,t}$, $AStkVol_{i,t-20,t-1}$, $AOptVol_{i,t-20,t-1}$, $ARange_{i,t-20,t-1}$, and $AVolatility_{i,t-20,t-1}$) at the 1st and 99th percent level to reduce the effect of outliers. The number of observations for *PJR* is lower than for the other variables in Table 1 because we exclude observations with low absolute values of *CAR*(-21,+2). The number of observations for abnormal shorting activity, abnormal intraday volatility and option volume are reduced because of the required match with CRSP using historical tickers and because of the availability of options volume data. The *PJR* has a mean value of 0.516, indicating that a substantial amount of total price discovery in the period from 1 month before the earnings announcement to 2 days after the earnings announcement takes place in the last 3 days. On average, we observe increased short selling activity and intraday volatility in the 20 days before earnings announcements relative to the 20-day period from day -50 to day -31. We observe slightly reduced trading activity in the options and stock market trading in the 20 days before the earnings announcement relative to the benchmark period from day -50 through day -31.

Several control variables used in our empirical models are averaged over the 20-day window (-50,-31): bid ask spread (*BIDASK*) measured as the closing ask minus the closing bid, divided by the mid-price; the natural log of the market capitalisation (*MV*) measured as the closing price times the number of shares outstanding and the natural log of the closing price. The natural log of the daily standard deviation of returns *(STDRET*) similarly is based on daily returns from day -50 through day -31. We also use the natural log of the number of analysts that provide

earnings estimates to IBES (*ANALYST*), and institutional ownership (*IO*) based on 13F. *PREVIOUS* is a dummy variable that equals one if the firm had a data breach between 2005 (the start of the Clearinghouse database) and the previous year, and *FORECAST* is a dummy variable that equals one if management provided earnings guidance in the 60-day window before the earnings announcement. *10KWords*, is the natural log of the total number of words in the most recent 10-K preceding the earnings announcement, and *10KMitiWords* is the natural log of the total number of mitigation words in the most recent 10-K preceding the earnings announcement.

In Table 1, Panel B we report the Pearson correlations between our variables of interest and main control variables. As might be expected, firms with longer cybersecurity-related discussions in their 10-Ks tend to have more mitigation words (the correlation between *NMitigation* and *NCyberWords* is 0.89). Correlations between *PJR*, our cybersecurity risk mitigation measures and several control variables are statistically significant, but tend to be low in magnitude (most correlations < |0.1|). The same applies to abnormal pre-announcement short selling and our measures of information asymmetry in the pre-announcement period. Not surprisingly, we observe relatively high correlations between abnormal short selling, option market volume and stock market volume. These last two variables are also highly correlated with the daily range in stock prices. Finally, high correlations are observed for the measures related to firm size such as bid-ask spread, market capitalization, number of analysts, institutional ownership and the number of words in 10-Ks.

(Insert Table 1 here)

4 Empirical method and results

4.1 Validation of Cybersecurity Mitigation Measures

To validate our two measures of cybersecurity risk mitigation (*CyberWords* and *Mitigation*), we conduct two tests to ensure that our measures: 1) plausibly reflect cybersecurity risk mitigation; and 2) perform better than alternative measures. In total, we consider four potential alternative measures of cybersecurity risk mitigation. The first is the negative tone of cybersecurity-related excerpts (*NegTone*) in a firm's 10-K, measured by the number of negative words to total words in these excerpts (Loughran and McDonald 2011). A more negative tone in the cybersecurity-related excerpts could reflect more risk mitigation, since negative wordings tend to be more cautionary in nature. The second alternative measure captures a firm's IT capital investment. *IT_Capexp* is natural log of one plus the total count of 10-K words related to IT software packages (Ashraf and Sunder 2018; Ashraf et al. 2020). These investments are likely to be associated with more investment in cybersecurity risk mitigation (Bose and Leung 2013; Chai et al. 2011; Im et al. 2001). The remaining alternative measures are based on the firm's corporate governance. Firms with better IT governance are more likely to be aware of cybersecurity issues and take mitigation actions (Chatterjee et al. 2001; Kwon et al. 2013). We explore two measures of IT governance: the number of IT executives in top management (*IT_Exec*) (Kwon et al. 2013); and the presence of an IT or risk board committee (*RiskComm*) (Higgs et al. 2016). Table 1, panel C reports descriptive statistics for these alternative cybersecurity risk mitigation variables. On average, 6% of words in the cyber excerpts are negative and the average number of IT executives in top management is 0.04. Consistent with Higgs et al. (2016), 10 percent of firms have a specific IT or Risk Committee.

The first validation test examines whether the measures are effective at predicting future data breaches in a firm. If cybersecurity mitigation is effective, firms that improve their cybersecurity risk mitigation score should be less likely to experience future cyber-related data breaches. To test this proposition, we estimate the following logistic model for each measure:

$$\begin{aligned} BREACH_{i,t+1} = a + \beta_1 CyberMiti_{i,t} + \beta_2 10KWords_{i,t} + \beta_3 10KMitiWords_{i,t} + \beta_4 PREVIOUS_{i,t} + \\ \beta_5 INDDIR_{i,t} + \beta_6 BRDSIZE_{i,t} + \beta_7 IO_{i,t} + \beta_8 MV_{i,t} + \beta_9 BM_{i,t} + \beta_{10} CAPEXP_{i,t} + \\ \beta_{11} INTANGIBLE_{i,t} + \beta_{12} RND_{i,t} + \beta_{13} ROA_{i,t} + \beta_{14} LEV_{i,t} + \beta_{15} SP500_{i,t} + \Sigma FirmFE + \Sigma YearFE + \varepsilon_{i,t} \end{aligned} \quad (4)$$

BREACH is an indicator variable coded one if the firm experiences a cybersecurity breach in the subsequent year (t+1) and zero otherwise. We obtain our sample of data breaches from the Privacy Rights Clearinghouse database (https://privacyrights.org/data-breaches). Following Kamiya et al. (2020), we only include breaches where a firm lost personal information subject to cyberattack notification laws. In addition, we exclude events classified as "CARDS", "PHYS" and "UNKN", which represent events related to the use of debit and credit card skimming devices at terminals, events where paper documents were lost or stolen, and events where it is not known how the information was exposed. We identify 130 breaches in our sample from this dataset. Our test variable of interest are the alternative cybersecurity risk mitigation measures (*CyberMiti*), which vary across models. We control for the number of words in a firm's 10-K (*10KWords*), the number of mitigation words in the 10-K (*10KMitiWords*), and for the existence of a prior cybersecurity breach (*PREVIOUS*). We also control for corporate governance characteristics including the proportion independent directors (*INDDIR*), board size (*BRDSIZE*) and institutional ownership (*IO*). We further control for firm financial characteristics, including firm size (*MV*), book to market ratio (*BM*), capital expenditures (*CAPEXP*), intangible assets (*INTANGIBLE*), research and development (*RND*), return on assets (*ROA*), leverage (*LEV*), and whether a firm is

included in the S&P 500 index (*SP500*). These variables are defined in Appendix B. Following Ashraf (2020), we include year and firm fixed effects to mitigate endogeneity concerns.

Results for this first validation test are provided in Table 2. Of the six cybersecurity risk mitigation measures, we find significant coefficients of *CyberMiti* in the expected direction when it is measured by *Mitigation* (at the one percent significance level), *CyberWords* (at the five percent significance level) and *NegTone* (at the five percent significance level). The results for the control variables show that firms that had a previous breach, firms with an increase in the proportion of intangible assets, greater capital expenditures, and with a higher number of mitigation words in the 10-K have a lower probability of a future breach. Firms with an increase in the number of words in their 10-K are more likely to have a future breach.

(Insert Table 2 here)

The second validation test examines the effect of cybersecurity breaches in a firm on its peer firms. Results in the literature indicate that when a firm is affected by a cybersecurity breach, peer firms are more likely to also be impacted in the future. For example, Ettredge and Richardson (2003) find that data breaches are correlated across firms in an industry. Ashraf (2020) finds that investors significantly increase downloads of peer firm's 10-K filings from EDGAR on the day that a firm discloses a breach. Kamiya et al. (2020, p. 27) conclude that "attacks are contagious to firms in the same industry." Based on these findings, we expect that when a firm's peer experiences a breach, the firm will be more likely to improve its cybersecurity risk management. A high quality measure of cybersecurity mitigation should exhibit a positive change subsequent to a breach at a peer firm.

Following Ashraf (2020), we identify industry peers using Hoberg-Phillips text-based network industry classifications (TNIC). TNIC categorizes firms as peers if they share a product

space (Hoberg and Phillips 2010, 2016). The database provides a set of peer firms that each have a TNIC score that ranges from 0 to 1, based upon the extent of product similarity. For the 130 data breaches in our sample, TNIC classifies 3,367 firms as 'peers' in the firm-years that the breaches occurred.

For each of our potential cybersecurity risk mitigation measures, we analyse the change in the measure between the year of the peer-breach and the next two years, using the following model:

$$\begin{aligned}\Delta CyberMiti_{i,t} = a + \beta_1 PEER_{i,t} + \beta_2 10KWords_{i,t} + \beta_3 10KMitiWords_{i,t} + \beta_4 PREVIOUS_{i,t} + \beta_5 INDDIR_{i,t} + \\ \beta_6 BRDSIZE_{i,t} + \beta_7 IO_{i,t} + \beta_8 MV_{i,t} + \beta_9 BM_{i,t} + \beta_{10} CAPEXP_{i,t} + \beta_{11} INTANGIBLE_{i,t} + \\ \beta_{12} RND_{i,t} + \beta_{13} ROA_{i,t} + \beta_{14} LEV_{i,t} + \beta_{15} SP500_{i,t} + \Sigma FF48FE + \Sigma YearFE + \varepsilon_{i,t}\end{aligned} \quad (5)$$

We measure *ΔCyberMiti* as either the one-year change in the cybersecurity risk mitigation measure from year t to year t+1 or the two-year change in the cybersecurity risk mitigation measure from year t to year t+2. Our variable of interest is *PEER,* which is an indicator variable coded one if the firm is classified as a peer firm of another firm that experienced a cybersecurity breach in year t and zero otherwise. The control variables used in this model specification are similar to the controls specified in equation (4). We include year and industry fixed effects and standard errors are clustered by firm. The industry classification is based on Fama and French 48 industry membership (Fama and French 1997)

The sample for the second validation test comprises 14,305 firm-years when we analyse the one-year change measures, and 11,011 firm-years when we analyse the two-year change measures. Table 3, panels A and B report the results for one-year change and the two-year change measures, respectively. We find that after a data breach at a peer firm, firms use, on average, 23 more *CyberWords* in their cybersecurity excerpts in the next year relative to firms that did not experience a peer breach, and 48 more cyber-security words after two years. Both coefficients of *PEER* are significant at the 5 percent level. Results for *Mitigation* are somewhat weaker. They

indicate an insignificant change in the number of mitigation words in the year after a peer breach occurs (Table 3, panel A), although on average, our sample firms significantly increase mitigation words by almost 1 after two years (significant at the 5 percent level, Table 3, panel B).[15] There is no evidence of a significant increase in any of our alternative measures of cybersecurity risk mitigation. We conclude that following a cybersecurity breach in a focal firm, peer firms increase their use of cybersecurity risk mitigation words and provide more extensive cybersecurity-related disclosures in their 10-Ks. Based on the result of our validation tests, we employ *CyberWords* and *Mitigation* as proxies for cybersecurity risk mitigation in the remainder of this study.

(Insert Table 3 here)

4.2 Information leakage and pre-announcement price discovery.

In this section we test our first hypothesis that privately informed trading by hackers is less likely if firms improve their cybersecurity risk mitigation, resulting in relatively slower discovery of new earnings information. Based on Weller (2018) we estimate the following model, with *PJR* as the dependent variable:

$$\begin{aligned} PJR_{i,t} = a + \beta_1 CyberMiti_{i,t} + \beta_2 10KWords_{i,t} + \beta_3 10KMitiWords_{i,t} + \beta_4 PREVIOUS_{i,t} + \beta_5 FORECAST_{i,t} + \\ \beta_6 ANALYST_{i,t} + \beta_7 IO_{i,t} + \beta_8 MV_{i,t} + \beta_9 BIDASK_{i,t} + \beta_{10} STDRET_{i,t} + \Sigma FirmFE + \Sigma YearQrtFE + \varepsilon_{i,t} \end{aligned} \quad (6)$$

Our variable of interest is the cybersecurity risk mitigation measure (*CyberMiti*), which is either *CyberWords* or *Mitigation*, depending on the model. We control for the natural log of the number of words in a firm's 10-K (*10KWords*), the natural log of the number of mitigation words in a firm's 10-K (*10KMitiWords*), whether the firm experienced a cybersecurity breach in the past (*PREVIOUS*), and whether management issued an earnings forecast (*FORECAST*) in the 60 trading days before the earnings announcement. The remaining control variables are based on

[15] This observation of a significant change in *Mitigation* words two years (but not one-year) after a peer breach occurs could be explained by firms requiring time to invest and implement better cybersecurity risk mitigation efforts.

Weller (2018) and include the natural log of the number of analysts covering stock *i* in quarter *t-1* (*ANALYST*) from I/B/E/S, and the institutional ownership ratio at the end of the preceding quarter from 13-F filings (*IO*). Weller (2018) also controls for the natural log of market capitalization (*MV*), bid ask spread (*BIDASK*), and the natural log of the standard deviation of daily returns (*STDRET*), all of which are computed over the period from event day -50 through event day -31. All models include year-quarter fixed effects, and firm fixed effects. Standard errors are two-way clustered by stock and quarter.

Table 4 presents the results from estimating equation (6). Column (1) presents results when we exclude observations where the absolute *CAR*(-21,+2) is smaller than the square root of 24 times the daily standard deviation of the returns in the period from day -50 to day -31 (the exclusion criterion used in Weller (2018)). Similar to Weller (2018), we find that the price jump at the earnings announcement is higher (or equivalently, that information leakage before the earnings announcement is lower) if firms are covered by a larger number of analysts and have higher institutional ownership. Further, the price jump at the earnings announcement decreases with quoted spread, and the price jump is lower for earnings announcements where firms issued an earnings-related management forecast.

For our test of Hypothesis 1, in column (1), the coefficient of *CyberMiti* using our first cybersecurity risk mitigation measure, *CyberWords,* is positive and significant at the five percent level. Regarding our second measure, *Mitigation*, the coefficient is positive and significant at the one percent level. The coefficient estimates suggest that a one-standard deviation increase in *CyberWords* increases the *PJR* by 0.014 (=1.76*0.008), which corresponds to a 2.8 percent increase in the fraction of the price discovery that occurs at the time of the public announcement. Similarly, a one-standard deviation increase in *Mitigation* increases the *PJR* by 0.025

(=1.18*0.021), which corresponds to a 4.9 percent increase in the fraction of the price discovery that occurs at the time of the public announcement. To put these numbers in perspective, note that issuing a management forecast decreases the PJR with about 0.050, and that a one standard deviation increase in institutional ownership increases the PJR with 0.031 (=0.11*0.287).[16]

Columns (2)-(6) employ alternative samples and an alternative PJR measure based on trading volume. The PJR-measure used in column (2), is based on a shorter pre-announcement window that starts on day -10 and, similar to the measure used in column (1), ends on day +2. The results in column (2) are similar to the results in column (1). Column (3) shows that when we drop all observations with an absolute *CAR*(-21,+2) below the median, we obtain results similar to column (1). Column (4) shows that when we drop all observations with an absolute *CAR*(-21,+2) below the 25th percentile, the adjusted R-squared drops and the results become considerably weaker because of the additional noise as a result of inclusion of PJRs with small denominators. The fifth column is based on the full sample. The adjusted R-squared drops even further and our cybersecurity risk mitigation measures become insignificant. Finally, we show in column (6) that our results for the full sample hold when using a volume-based PJR measure, *PJR_Volume*.[17] Analogous to equation (1), this volume-based PJR measure for each earnings announcement is calculated as the cumulative volume in the 3-day window from day 0 to day +2 divided by the cumulative volume over a longer window starting on day -21 and ending day +2. For *PJR_Volume*, there is no small denominator problem and we can use the full sample. The coefficients for our cybersecurity risk mitigation proxies, *CyberWords* and *Mitigation,* are both positive and

[16] In an untabulated univariate analysis following Weller (2018), we examine the difference in return for perfect foresight portfolios based upon the outer quintiles of *Mitigation*. We find that the pre-announcement returns are significantly higher for the low cybersecurity risk mitigation portfolio, with the difference accelerating starting from approximately 20 days prior to the earnings announcement and reaching more than 7 percent on the day before the earnings announcement. This difference is statistically significant, t=6.16, p < 0.0001.

[17] We thank one of the anonymous referees for this suggestion.

significant at the one percent level, which indicates that firms that improve their cybersecurity policies on average have relative less trading taking place in the pre-announcement period.

(Insert Table 4 here)

Overall, based on the evidence in Table 4, we conclude that firms with improved cybersecurity risk mitigation have less information leakage in pre-announcement period.

4.3 Cybersecurity risk mitigation and short selling

In this section we test our second hypothesis, that firms with improved cybersecurity risk mitigation have less informative short selling in the days before earnings announcements. We obtain data on short sales transactions from the Financial Industry Regulatory Authority (FINRA) website.[18] Beginning in August 2009, FINRA provides data of short sale transactions that include transaction times, prices, and sizes for all short sales of National Market System stocks.

To test the impact of cybersecurity risk mitigation on the ability of short sellers to predict earnings surprises, we specify the following model:

$$\begin{aligned} RSURPRISE_{i,t} = {} & \alpha + \beta_1 CyberMiti_{i,t} + \beta_2 ASHVOL(-20,-1)_{i,t} + \beta_3 CyberMiti_{i,t} \times ASHVOL(-20,-1)_{i,t} + \\ & \beta_4 10KWords_{i,t} + \beta_5 10KMitiWords_{i,t} + \beta_6 PREVIOUS_{i,t} + \beta_7 FORECAST_{i,t} + \beta_8 MV_{i,t} + \\ & \beta_9 BM_{i,t} + \beta_{10} PRC_{i,t} + \beta_{11} BIDASK_{i,t} + \beta_{12} TURN_{i,t} + \beta_{13} STDRET_{i,t} + \beta_{14} CAR(-50,-21)_{i,t} + \\ & \beta_{15} CAR\ (-20,-1)_{i,t} + \Sigma FirmFE + \Sigma YearQrt + \varepsilon_{i,t} \end{aligned} \quad (7)$$

where *RSURPRISE* is the quarterly rank decile of the earnings surprise, with earnings surprise defined as the difference between the actual earnings per share and the most recent average earnings per share forecast from I/B/E/S, scaled by the absolute value of this most recent average earnings per share forecast. As before, *CyberMiti* is either *CyberWords* or *Mitigation*.

[18] The FINRA short transactions data have been used in prior studies (e.g., Berkman and Eugster 2017; Jain et al. 2013; Jain et al. 2012). For more information on the short sales transactions on FINRA, see http://www.finra.org/sites/default/files/NoticeDocument/p120044.pdf .

$ASHVOL(-20,-1)$ is the average daily abnormal short volume over a 20-day window before the earnings announcement. All other control variables have been defined before, with the exception of *TURN,* which is the number of shares traded during the day as a proportion of shares outstanding, also averaged over event day -50 through -31, *PRC*, which is the natural log of the closing share price, and the cumulative abnormal returns over event days -50 to -21, *CAR*(-50,-21)*,* and over the 20-day window, *CAR*(-20,-1). Year-quarter fixed effects and firm fixed effects are included, and standard errors are clustered at both the firm and quarter levels.

The key independent variable in the regression is the interaction term for *CyberWords (Mitigation)* and $ASHVOL(-20,-1)$. Our main hypothesis is that pre-announcement informed short sales are less likely for firms that improve their cybersecurity risk mitigation. If informed short selling prior to the announcement of earnings surprises is less likely when firms improve their cybersecurity risk mitigation, we expect a significant positive coefficient of the interaction term between our cybersecurity risk mitigation measures and pre-announcement short selling. That is, the relation between earnings surprise and pre-announcement short selling is less negative when firms have relatively higher values of *CyberWords* or *Mitigation*.

The results of this analysis are presented in panel A of Table 5. Results in column (1) show that earnings surprises are negatively related to firm size and positively associated with the stock price, book-to-market ratio and the pre-announcement stock returns for both pre-announcement windows (-50,-21) and (-20,-1). In line with our prediction, the interaction term between *ASHVOL* and *CyberWords* is positive and significant at the one percent level. Likewise, the interaction term between *ASHVOL* and *Mitigation* is positive and significant at the five percent level. This result indicates that when firms have improved cybersecurity risk mitigation, the relation between earnings surprise and pre-earnings announcement short selling is less negative.

Columns (2) - (4) in Table 5 present the results of several alternative measures of earnings surprise. The dependent variable in column (2) is the quarterly rank decile of the earnings surprise, with earnings surprise defined as the difference between the actual earnings per share and the most recent average earnings per share forecast from I/B/E/S, but now scaled by the stock price 10 days before the earnings announcement. Column (3) presents results when the earnings surprise measure is calculated as the difference between the actual earnings per share and the most recent average earnings per share estimate across analysts, scaled by the standard deviation across analyst estimates. Finally, column (4) presents results using the median earnings per share forecast across analysts instead of the mean earnings per share forecast to calculate the earnings surprise measure. The results in columns (2) - (4) are all consistent with the specification in column (1) and consistently show a positive and significant association between the quarterly earnings surprise rank decile and the interaction variable between our cybersecurity risk mitigation variables (*CyberWords* and *Mitigation*) and abnormal short selling in the month before the earnings announcement.

To gain further insight in the timing of the privately informed trading by short sellers, panel B of Table 5, reports the results that examine short sales in one-week windows [weeks(-1,-6)] leading up to earnings announcement. The coefficients of our test variable of interest, the interaction between *ASHVOL* and our cybersecurity risk mitigation measures, are generally significant at the 10 percent level or better and in the expected positive direction for week -1 through to week -4 (columns 1-4). In weeks -5 and -6, the interactions between *ASHVOL* and our cybersecurity risk mitigation measures are substantially smaller and no longer significant. These results suggest that the window over which digital insiders trade extends approximately over one month prior to earnings announcements. We note that in addition to short selling driven by the

activity of informed traders, other (uninformed) market participants may take cues from changes in trading volume and trade "with the wind" (Van Kervel and Menkveld 2019). Such trades will exacerbate the impact of the illegally-obtained information on trade volumes and price formation.

(Insert Table 5 here)

4.4 Cybersecurity and the probability of informed trading before earnings announcements

In this section, we test the third hypothesis, which predicts that firms that improve their cybersecurity have lower abnormal stock and option volume, and lower intraday volatility in the pre-earnings announcement period. As discussed before, we select the 50 trading days before each earnings announcement for our sample stocks. For each earnings announcement, we calculate abnormal values for each measure for day -20 through day -1, using the mean and standard deviation based on the 21 trading days from day -50 to day -30. In our primary analyses, we use four measures of informed trading: (1) Stock trading volume (*StkVol*), (2) Option trading volume (*OptVol*), Price range (*Range*) and Intraday volatility (*Volatility*). For each variable, we calculate the average of *A_ASINFO* over the 20-day window before the earnings announcement and estimate the following model using this average abnormal information asymmetry measure as dependent variable:

$$\begin{aligned} A\_ASINFO_{1-4,i,q,t} = a + \beta_1 CyberMiti_{i,q} + \beta_2 10KWords_{i,q} + \beta_3 10KMitiWords_{i,q} + \\ \beta_4 PREVIOUS_{i,q} + \beta_5 FORECAST_{i,q} + \beta_6 PRC_{i,q} + \beta_7 MV_{i,q} + \\ \beta_8 TURN_{i,q} + \beta_9 STDRET_{i,q} + \Sigma FirmFE + \Sigma YearQrtFE + \varepsilon_{i,q,t} \end{aligned} \quad (8)$$

CyberMiti is either *CyberWords* or *Mitigation*. Model 8 includes standard control variables in regression models of measures of information asymmetry (e.g., Kacperczyk and Pagnotta 2019; Stoll 1989), and all variables have been defined before.

In addition to the 20-day window leading up to the earnings announcement, we also estimate the model for each of the 6 weeks leading up to the earnings announcement and the week after the earnings announcement. Based on the results in the previous section, we expect to see evidence of increased information asymmetry in the month before earnings announcements, peaking around week -3.

Panel A of Table 6 reports the results where equation (8) is estimated using a 20-day window preceding quarterly earnings announcements. Column (1) presents the results with abnormal stock trading volume (*AStkVol*) as the measure of pre-announcement information asymmetry. Column (2) presents the results with abnormal option trading volume (*AOptVol*) as the information asymmetry measure. Column (3) presents the results with abnormal price range (*ARange*) as the information asymmetry measure. Column (4) presents the results with abnormal intraday volatility (*AVolatility*) as the information asymmetry measure. The coefficients for *CyberWords* and *Mitigation* are significant in the expected negative direction at the five percent level or better for all measures of information asymmetry. Collectively, these results indicate that when a firm has better cybersecurity risk mitigation policies, it is expected to have lower abnormal stock and option volume in the pre-earnings announcement period and reduced intraday volatility.

Panel B of Table 6 reports the coefficients for *CyberMiti* in equation (8) for various windows preceding or following the quarterly earnings announcements. In this panel, we also present results using other liquidity measures that are commonly used in the accounting and finance literatures. For our four primary information asymmetry measures, (*AStkVol*, *AOptVol*, *ARange* and *AVolatility*) we find that the coefficient of *Mitigation* is generally significant in the expected negative direction from weeks -1 through -5 (columns 2 – 6), peaking in week -3 or -4

(columns 4 and 5). For *CyberWords*, the results are somewhat weaker but display the same general pattern.

In addition to our primary measures of informed trading, we include analysis of measures of information asymmetry used in prior research. Focusing on the 20-day window in the first column in Panel B, we see that the effective spread (*ES*), the quoted spread (*QS*), the price impact (*PI*) measure and Kyle's lambda (*Lambda*) are all significant with the expected sign. The results for the realized spread (*RS*), the Amihud (*AMI*) measure and order imbalance (*OI*) are insignificant, and in some cases have the opposite sign to what is expected. These results are similar to Akey et al. (2020), who also report a lack of significance for the order imbalance measure and the Amihud illiquidity measure, and find evidence that an increased probability of trading by informed traders increases spread-based measures.

Overall, we find that our primary measures of the probability of informed trading are significantly negatively related to a firm's cybersecurity risk mitigation. We also find that the impact of cybersecurity risk mitigation on our measures of information asymmetry peaks around the third and fourth week preceding quarterly earnings announcements.

(Insert Table 6 here)

5 Conclusion

There is growing awareness that hackers target corporations to obtain non-public corporate information for illegal trading. Spurred by growing concerns about hacking activities, this study examines whether a firm's cybersecurity risk mitigation affects the extent to which its private information is traded on and is reflected in prices before earnings announcements. We develop and employ two measures of cybersecurity risk mitigation for our study. Our first measure captures a firm's cybersecurity risk mitigation by simply counting words in cyber-related disclosures in their

10-Ks. Our second measure captures a firm's cybersecurity risk mitigation by only counting mitigation words in the cybersecurity-related excerpts in a firm's 10-K. We validate these measures by showing they are predictive of future cybersecurity breaches in a firm, and showing a significant increase in these measures in reaction to cybersecurity breaches at peer firms. Our measures perform better in these tests than alternative measures of cybersecurity risk mitigation from the literature.

In our main tests, we first provide evidence that firms with lower cybersecurity risk mitigation experience smaller price changes at the time of the earnings announcement (i.e., a low price jump ratio), indicating that a greater amount of information is reflected in stock prices before this information is publicly revealed. Further tests indicate that pre-announcement trading by short sellers is more predictive of earnings announcement surprises for firms with low cybersecurity risk mitigation scores. Finally, we demonstrate that firms with lower cybersecurity risk mitigation scores experience a larger increase in measures of informed trading in the weeks before earnings announcements. Collectively, these findings indicate that firms with better cybersecurity risk mitigation are associated with a lower extent to which private information is traded on and reflected in prices before publication of new earnings information. This evidence is consistent with leakage of information through hacking prior to earnings announcements. Overall, our findings suggest that weak cybersecurity risk mitigation increases opportunities for acquisition of private information by hackers, resulting in an increased probability of informed trading in the pre-earnings announcement period.

Our results should be of interest to a variety of firm stakeholders. We find evidence that cybersecurity risk mitigation disclosures help to explain variation in price formation across firms. We provide evidence of tangible benefits from cybersecurity risk mitigation, which should be of

interest to regulators, helping them in the trade-off between the costs and benefits of specific forms of cyber mitigation and related disclosure requirements. Our results indicate that investors can benefit from evaluating the extant cybersecurity risk mitigation policies of listed companies so that they can better understand price formation for those firms. Finally, given an improved understanding of the consequences for shareholders of having more or less cybersecurity mitigation, management can better understand the tradeoffs they face when evaluating their investments in cybersecurity mitigation activities.

References

Ahern, K. R. 2020. Do proxies for informed trading measure informed trading? Evidence from illegal insider trades. *The Review of Asset Pricing Studies* 10 (3):397-440.

Akey, P., V. Grégoire, and C. Martineau. 2020. Price Revelation from Insider Trading: Evidence from Hacked Earnings News. *Available at SSRN 3365024*.

Alford, A., J. Jones, R. Leftwich, and M. Zmijewski. 1993. The relative informativeness of accounting disclosures in different countries. *Journal of Accounting Research* 31:183-223.

Amihud, Y. 2002. Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets* 5 (1):31-56.

Amihud, Y., and H. Mendelson. 1980. Dealership market: Market-making with inventory. *Journal of Financial Economics* 8 (1):31-53.

Amir, E., S. Levi, and T. Livne. 2018. Do firms underreport information on cyber-attacks? Evidence from capital markets. *Review of Accounting Studies* 23:1177–1206.

Ashraf, M. 2020. The role of market forces and regulation in disclosure: Evidence from cyber risk factors. Working paper.

Ashraf, M., P. N. Michas, and D. Russomanno. 2020. The Impact of Audit Committee Information Technology Expertise on the Reliability and Timeliness of Financial Reporting. *The Accounting Review* 95 (5):23-56.

Ashraf, M., and J. Sunder. 2018. Consumer protection regulation and the cost of equity: Evidence from data breach disclosure laws. *Available at SSRN 3308551*.

Ball, R., and P. Brown. 1968. An Empirical Evaluation of Accounting Income Numbers. *Journal of Accounting Research* 6 (2):159-178.

Ball, R., and L. Shivakumar. 2008. How much new information is there in earnings? *Journal of Accounting Research* 46 (5):975-1016.

Berkman, H., and M. Eugster. 2017. Short on drugs: Short selling during the drug development process. *Journal of Financial Markets* 33:102-123.

Berkman, H., J. Jona, G. Lee, and N. S. Soderstrom. 2018. Cybersecurity Awareness and Market Valuations. *Journal of Accounting and Public Policy* 37 (6):508-526.

Berkman, H., M. D. McKenzie, and P. Verwijmeren. 2016. Hole in the wall: Informed short selling ahead of private placements. *Review of Finance* 21 (3):1047-1091.

Bessembinder, H. 2003. Issues in assessing trade execution costs. *Journal of Financial Markets* 6 (3):233-257.

Bessembinder, H., and H. M. Kaufman. 1997. A comparison of trade execution costs for NYSE and NASDAQ-listed stocks. *Journal of Financial and Quantitative Analysis* 32 (3):287-310.

Bianchi, D., and O. K. Tosun. 2019. Cyber attacks and stock market activity. *Available at SSRN 3190454*.

Bose, I., and A. C. M. Leung. 2013. The impact of adoption of identity theft countermeasures on firm value. *Decision Support Systems* 55 (3):753-763.

Brogaard, J., B. Hagströmer, L. Nordén, and R. Riordan. 2015. Trading fast and slow: Colocation and liquidity. *The Review of Financial Studies* 28 (12):3407-3443.

Bushee, B. J., J. E. Core, W. Guay, and S. J. Hamm. 2010. The role of the business press as an information intermediary. *Journal of Accounting Research* 48 (1):1-19.

Butler, M., A. Kraft, and I. S. Weiss. 2007. The effect of reporting frequency on the timeliness of earnings: The cases of voluntary and mandatory interim reports. *Journal of Accounting and Economics* 43 (2-3):181-217.

Cao, J., A. Goyal, S. Ke, and X. Zhan. 2020. Options Trading and Stock Price Informativeness. *Swiss Finance Institute Research Paper* (19-74).

Carson, J. 2017. The evolution of the digital insider trader. *Computer Fraud & Security* (8):12-15.

Cavusoglu, H., B. Mishra, and S. Raghunathan. 2004. The effect of internet security breach announcements on market value: Capital market reactions for breached firms and internet security developers. *International Journal of Electronic Commerce* 9 (1):70-104.

Chai, S., M. Kim, and H. R. Rao. 2011. Firms' information security investment decisions: Stock market evidence of investors' behavior. *Decision Support Systems* 50 (4):651-661.

Chakrabarty, B., and A. Shkilko. 2013. Information transfers and learning in financial markets: Evidence from short selling around insider sales. *Journal of Banking & Finance* 37 (5):1560-1572.

Chatterjee, D., V. J. Richardson, and R. W. Zmud. 2001. Examining the shareholder wealth effects of announcements of newly created CIO positions. *MIS Quarterly*:43-70.

Chen, Y., B. Kelly, and W. Wu. 2020. Sophisticated investors and market efficiency: Evidence from a natural experiment. *Journal of Financial Economics*.

Christophe, S. E., M. G. Ferri, and J. J. Angel. 2004. Short-selling prior to earnings announcements. *The Journal of Finance* 59 (4):1845-1876.

Christophe, S. E., M. G. Ferri, and J. Hsieh. 2010. Informed trading before analyst downgrades: Evidence from short sellers. *Journal of Financial Economics* 95 (1):85-106.

Collin-Dufresne, P., and V. Fos. 2015. Do prices reveal the presence of informed trading? *The Journal of Finance* 70 (4):1555-1582.

Easley, D., and M. O'Hara. 1987. Price, trade size, and information in securities markets. *Journal of Financial Economics* 19 (1):69-90.

Engelberg, J. E., A. V. Reed, and M. C. Ringgenberg. 2012. How are shorts informed?: Short sellers, news, and information processing. *Journal of Financial Economics* 105 (2):260-278.

Ettredge, M. L., and V. J. Richardson. 2003. Information transfer among internet firms: the case of hacker attacks. *Journal of Information Systems* 17 (2):71-82.

Fama, E. F., and K. R. French. 1992. The cross-section of expected stock returns. *The Journal of Finance* 47 (2):427-465.

———. 1997. Industry costs of equity. *Journal of Financial Economics* 43 (2):153-193.

Ferraro, M. F. 2014. Groundbreaking or Broken; An Analysis of SEC Cybersecurity Disclosure Guidance, Its Effectiveness, and Implications. *Albany Law Review* 77:297-347.

Fischer, E. A. 2014. Federal Laws Relating to Cybersecurity: Overview of Major Issues, Current Laws, and Proposed Legislation: Congressional Research Service.

Glosten, L. R., and L. E. Harris. 1988. Estimating the components of the bid/ask spread. *Journal of Financial Economics* 21 (1):123-142.

Glosten, L. R., and P. R. Milgrom. 1985. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics* 14 (1):71-100.

Gordon, L. A., M. P. Loeb, W. Lucyshyn, and L. Zhou. 2015. Increasing cybersecurity investments in private sector firms. *Journal of Cybersecurity* 1 (1):3-17.

Gordon, L. A., M. P. Loeb, and T. Sohail. 2010. Market value of voluntary disclosures concerning information security. *MIS Quarterly*:567-594.

Gordon, L. A., M. P. Loeb, and L. Zhou. 2011. The impact of information security breaches: Has there been a downward shift in costs? *Journal of Computer Security* 19 (1):33-56.

Hamm, K. M. *Cybersecurity: Where We Are; What More Can be Done? A Call for Auditors to Lean In*. PCAOB 2019 [cited 05/11/2019. Available from https://pcaobus.org/News/Speech/Pages/hamm-cybersecurity-where-we-are-what-more-can-be-done.aspx.

Hasbrouck, J. 1991. Measuring the information content of stock trades. *The Journal of Finance* 46 (1):179-207.

———. 2009. Trading costs and returns for US equities: Estimating effective costs from daily data. *The Journal of Finance* 64 (3):1445-1477.

Higgs, J. L., R. E. Pinsker, T. J. Smith, and G. R. Young. 2016. The Relationship between Board-Level Technology Committees and Reported Security Breaches. *Journal of Information Systems* 30 (3):79-98.

Hilary, G., B. Segal, and M. H. Zhang. 2016. Cyber-Risk Disclosure: Who Cares? Georgetown McDonough School of Business Research Paper No. 2852519, Available at SSRN: https://ssrn.com/abstract=2852519 or http://dx.doi.org/10.2139/ssrn.2852519.

Ho, T., and H. R. Stoll. 1981. Optimal dealer pricing under transactions and return uncertainty. *Journal of Financial Economics* 9 (1):47-73.

Ho, T. S., and H. R. Stoll. 1983. The dynamics of dealer markets under competition. *The Journal of Finance* 38 (4):1053-1074.

Hoberg, G., and G. Phillips. 2010. Product market synergies and competition in mergers and acquisitions: A text-based analysis. *The Review of Financial Studies* 23 (10):3773-3811.

———. 2016. Text-based network industries and endogenous product differentiation. *Journal of Political Economy* 124 (5):1423-1465.

Holden, C. W., S. Jacobsen, and A. Subrahmanyam. 2014. The empirical analysis of liquidity. *Foundations and Trends in Finance* 8 (4):263-365.

Hovav, A., and J. D'arcy. 2005. Capital market reaction to defective IT products: The case of computer viruses. *Computers & Security* 24 (5):409-424.

Huang, H. H., and C. Wang. 2020. Do Banks Price Firms' Data Breaches? . *The Accounting Review* forthcoming.

Huang, Q., and T. R. Skantz. 2016. The informativeness of pro forma and street earnings: an examination of information asymmetry around earnings announcements. *Review of Accounting Studies* 21 (1):198-250.

Huang, R. D., and H. R. Stoll. 1996. Dealer versus auction markets: A paired comparison of execution costs on NASDAQ and the NYSE. *Journal of Financial Economics* 41 (3):313-357.

Im, K. S., K. E. Dow, and V. Grover. 2001. A reexamination of IT investment and the market value of the firm—An event study methodology. *Information Systems Research* 12 (1):103-117.

Jain, A., P. K. Jain, T. H. McInish, and M. McKenzie. 2013. Worldwide reach of short selling regulations. *Journal of Financial Economics* 109 (1):177-197.

Jain, C., P. Jain, and T. H. McInish. 2012. Short selling: the impact of SEC rule 201 of 2010. *Financial Review* 47 (1):37-64.

Kacperczyk, M., and E. S. Pagnotta. 2019. Chasing private information. *The Review of Financial Studies* 32 (12):4997-5047.

Kamiya, S., J.-K. Kang, J. Kim, A. Milidonis, and R. M. Stulz. 2018. What is the Impact of Successful Cyberattacks on Target Firms?: National Bureau of Economic Research.

———. 2020. Risk management, firm reputation, and the impact of successful cyberattacks on target firms. *Journal of Financial Economics*.

Karpoff, J. M., and X. Lou. 2010. Short sellers and financial misconduct. *The Journal of Finance* 65 (5):1879-1913.

Khan, M., and H. Lu. 2013. Do short sellers front-run insider sales? *The Accounting Review* 88 (5):1743-1768.

Kwon, J., J. R. Ulmer, and T. Wang. 2013. The association between top management involvement and compensation and information security breaches. *Journal of Information Systems* 27 (1):219-236.

Kyle, A. S. 1985. Continuous auctions and insider trading. *Econometrica: Journal of the Econometric Society*:1315-1335.

Lee, C., and M. J. Ready. 1991. Inferring trade direction from intraday data. *The Journal of Finance* 46 (2):733-746.

Lin, J.-C., G. C. Sanger, and G. G. Booth. 1995. Trade size and components of the bid-ask spread. *The Review of Financial Studies* 8 (4):1153-1183.

Loughran, T., and B. McDonald. 2011. When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance* 66 (1):35-65.

Massoud, N., D. Nandy, A. Saunders, and K. Song. 2011. Do hedge funds trade on private information? Evidence from syndicated lending and short-selling. *Journal of Financial Economics* 99 (3):477-499.

McNichols, M. 1984. *The anticipation of earnings in securities markets*: University of California, Los Angeles--Management.

Mitts, J., and E. L. Talley. 2019. Informed trading and cybersecurity breaches. *Harv. Bus. L. Rev.*:1-51.

Modi, S. B., M. A. Wiles, and S. Mishra. 2015. Shareholder value implications of service failures in triads: The case of customer information security breaches. *Journal of Operations Management* 35:21-39.

Roll, R. 1984. A simple implicit measure of the effective bid-ask spread in an efficient market. *The Journal of Finance* 39 (4):1127-1139.

SEC. 2011. CF Disclosure Guidance: Topic No. 2. Available at: https://www.sec.gov/divisions/corpfin/guidance/cfguidance-topic2.htm.

———. 2018. Commission Statement and Guidance on Public Company Cybersecurity Disclosures, Available at: https://www.sec.gov/rules/interp/2018/33-10459.pdf.

Selznick, L. F., and C. LaMacchia. 2016. Cybersecurity: Should the SEC Be Sticking Its Nose under This Tent. *Journal of Law, Technology & Policy* 16 (1):35-70.

Spanos, G., and L. Angelis. 2016. The impact of information security events to the stock market: A systematic literature review. *Computers & Security* 58:216-229.

Steinbart, P. J., R. L. Raschke, G. Gal, and W. N. Dilla. 2018. The influence of a good relationship between the internal audit and information security functions on information security outcomes. *Accounting, Organizations and Society* 71:15-29.

Stoll, H. R. 1989. Inferring the components of the bid-ask spread: theory and empirical tests. *The Journal of Finance* 44 (1):115-134.

Telang, R., and S. Wattal. 2007. An empirical analysis of the impact of software vulnerability announcements on firm stock price. *IEEE Transactions on Software Engineering* (8):544-557.

Van Kervel, V., and A. J. Menkveld. 2019. High-frequency trading around large institutional orders. *The Journal of Finance* 74 (3):1091-1137.

Wang, T., K. N. Kannan, and J. R. Ulmer. 2013. The association between the disclosure and the realization of information security risk factors. *Information Systems Research* 24 (2):201-218.

Weller, B. M. 2018. Does algorithmic trading reduce information acquisition? *The Review of Financial Studies* 31 (6):2184-2226.

Westland, J. C. 2018. The Information Content of Sarbanes-Oxley in Predicting Security Breaches. *Computers & Security* 90:1-20.

Yayla, A. A., and Q. Hu. 2011. The impact of information security events on the stock value of firms: The effect of contingency factors. *Journal of Information Technology* 26 (1):60-77.

Appendix A

A1. Cybersecurity Risk Mitigation Wordlist

actionable
address the risk
analysis of our security
analysis of our technologies
ciso
compliance
confidentiality agreement
consult
data management
detect
develop new polic(y|ies)
encrypt
enhancement
expert
hardened firewall
implement
independent analysis
independent review
insur
mitigat
monitor
monitoring solution
personnel
polic(y|ies) review
policies procedures and controls
policy solution
pre(| |\\-)emptive
predict
prevent
privacy and security management
privacy program
project(i|e)
protect
protocol
remedi
risk and fraud management
security analysis
security brief
security capabil
security enhancement
security process
security program
security protocol
security review
security solution
security tool
self(|\\-)regulatory
standards review
strategy
tak(e|ing) step
train
validat
verif
voluntary self(|\\-)disclosure
vulnerability assessment
vulnerability management

A2. Examples of Cyber Excerpts Containing Risk Mitigation Themes

1. Ellie Mae (FY 2015, reporting date: Feb 25, 2016)

16 cybersecurity-related excerpts were identified with a total *Mitigation* score of 51. Below is an example for 1 out of the 16 excerpts.

"All sensitive data transmitted over public networks is **encrypt**ed using industry standard **encrypt**ion **protocol**s in order to **protect** sensitive data against third-party disclosure in transit. Servers and network components are secured with access control mechanisms and **protect**ed by **hardened firewall**s, virus **protect**ion, and intrusion **prevent**ion/**detect**ion systems. Security services are **monitor**ed and updated in order to address emerging vulnerabilities. Even with our current security **monitor**ing and **detect**ion systems, we cannot guarantee that our security measures will **prevent** security breaches. We are committing significant resources to **protect** against and remedy any potential security breaches and their consequences and intend to keep doing so in the future. New threats and vulnerabilities are identified frequently and there are often time lags before our vendors deploy **mitigat**ions. In 2015 we made substantial investment in our network security infrastructure, including headcount and third party tools and systems. In 2016 and beyond we will continue to make substantial investments in our network security infrastructure to **protect** the confidentiality of the information stored in our data centers."

For this excerpt *NMitigation* score: 16

2. Capella Education Company (FY 2015, reporting date: Feb 18, 2016)

4 cybersecurity-related excerpts were identified with a total *Mitigation* score of 15. Below is an example for 1 out of the 4 excerpts.

"Capella has an information **security program** that includes leadership, tools, processes, and **train**ing. To **protect** our information assets, Capella's information security practices are designed to reduce information security and IT risks, respond to incidents, establish appropriate standards and controls, and establish, **implement**, and maintain information security policies and procedures. These practices include an education and **train**ing program on information security and privacy matters for employees and external stakeholders."

For this excerpt *NMitigation* score: 5

3. Fossil Group (FY 2016, reporting date: Feb 29, 2016)

6 cybersecurity-related excerpts were identified with a total *Mitigation* score of 7. Below is an example for 1 out of the 4 excerpts.

"We may experience operational problems with our information systems as a result of system failures, viruses, computer "hackers" or other causes. Any material disruption or slowdown of our systems could cause information, including data related to customer orders, to be lost or delayed which could result in delays in the delivery of merchandise to our stores and customers or lost sales, which could reduce demand for our merchandise and cause our sales to decline. Moreover, the failure to maintain, or a disruption in, financial and management control systems could have a material adverse effect on our ability to respond to trends in our target markets, market our products and meet our customers' requirements."

For this excerpt *NMitigation* score: 0

Appendix B: Variable Definitions

Variable	Definition	Source
Cybersecurity risk mitigation measures		
NCyberWords	Number of words in all cyber-related excerpts in a firm's disclosures in the entire 10-K filing for a given year.	10-Ks
CyberWords	Natural log of one plus *NCyberWords.*	
NMitigation	Number of cyber mitigation words or word combinations from Appendix A1 in all cyber-related excerpts in a firm's disclosures in the entire 10-K filing for a given year.	10-Ks
Mitigation	Natural log of one plus *NMitigation.*	
Alternative cybersecurity risk mitigation measures		
IT_Exec	Number of IT executives in top management.	BoardEx
RiskComm	Indicator variable coded one if a firm has an IT or Risk committee, zero otherwise.	BoardEx
IT_Capexp	The natural log of one plus the total count of 10-K words related to IT software packages (Ashraf and Sunder 2018; Ashraf et al. 2020).	10-Ks
NegTone	Number of negative words in cybersecurity excerpts in a firm's 10-K divided by total number of words in cybersecurity excerpts in a firm's 10-K.	10-Ks
PJR, Earnings surprise, short selling and information asymmetry measures		
PJR	Price Jump Ratio, computed as abnormal returns relative to a Fama and French (1992) three-factor model using daily returns over a 365-calendar day window ending 90 days before the earnings announcement. The pre-announcement window starts on day -21 and ends 2 trading days after the earnings announcement. The announcement return window starts on day 0 and ends 2 days after the earnings announcement.	CRSP
RSURPRISE	The quarterly rank decile of the earnings surprise, where earnings surprise is defined as the difference between the actual earnings per share and the most recent average earnings per share forecast from I/B/E/S, scaled by the absolute value of this most recent average earnings per share forecast.	I/B/E/S
$ASHVOL(-20,-1)$	Average daily abnormal short volume over the four weeks week before the earnings announcement. We use the mean and standard deviation of *SHVOL* over the 20 trading days from day -50 to day -31 to standardize the daily measures in the 20 days before the earnings announcement. We then calculate the average daily abnormal values for the 20 days before the earnings announcement to obtain $ASHVOL_{i,t\text{-}20,t\text{-}1.}$	FINRA
AStkVol(-20,-1)	Abnormal stock trading volume. Stock trading volume (*StkVol*) is measured as the natural log of the number of traded shares. We use the mean and standard deviation of *StkVol* over the 20 trading days from day -50 to day -31 to standardize the daily measures in the 20 days before the earnings announcement. We then calculate the average daily abnormal values for the 20 days before the earnings announcement to obtain $AStkVol_{i,t\text{-}20,t\text{-}1.}$	CRSP
AOptVol(-20,-1)	Abnormal option trading volume. Option trading volume (*OptVol*) is measured as the natural log of total volume in the call and put options of all strikes and all maturities. We use the mean and standard deviation of *OptVol* over the 20 trading days from day -50 to day -31 to standardize the daily measures in the 20 days before the earnings announcement. We then calculate the average daily abnormal values for the 20 days before the earnings announcement to obtain $AOptVol_{i,t\text{-}20,t\text{-}1.}$	OptionMetrics
ARange(-20,-1)	Abnormal price range. Price range (*Range*) is measured as the difference between the highest trade price minus the lowest trading price. We use the mean and standard deviation of *Range* over the 20 trading days from day -50 to day -31 to standardize the daily measures in the 20 days before the earnings announcement. We then calculate the average daily abnormal values for the 20 days before the earnings announcement to obtain $ARange_{i,t\text{-}20,t\text{-}1.}$	DTAQ
AVolatility(-20,-1)	Abnormal intraday volatility. Intraday volatility (*Volatility*) is the intraday trade-based volatility during trading hours. We use the mean and standard deviation of *Volatility* over the 20 trading days from day -50 to day -31 to standardize the daily measures in the 20 days before the earnings announcement. We then calculate the average daily abnormal values for the 20 days before the earnings announcement to obtain $AVolatility_{i,t\text{-}20,t\text{-}1.}$	DTAQ
Firm controls		
10KWords	Natural log of the total number of words in the most recent 10-K preceding the earnings announcement.	10-Ks

10KMitiWords	Natural log of the total number of mitigation words in the most recent 10-K preceding the earnings announcement.	10-Ks
PREVIOUS	Indicator variable coded one if the firm experienced a cybersecurity breach in the prior year or earlier, zero otherwise. We follow Kamiya et al. (2020) and only include breaches where a firm lost personal information subject to cyberattack notification laws. We exclude events classified as "CARDS", "PHYS" and "UNKN" in the Privacy Rights Clearinghouse database.	Privacy Rights Clearinghouse
FORECAST	Indicator variable coded one if management issued earnings forecast, zero otherwise.	I/B/E/S
ANALYST	Natural log of number of analysts covering the stock.	I/B/E/S
INDDIR	Percentage of independent directors on the board.	BoardEx
BRDSIZE	The natural log of the number of directors on the board.	BoardEx
IO	Institutional ownership ratio at the end of the preceding quarter.	Thomson Reuters
MV	The natural log of the stock's market capitalization measured as closing price times the number of shares outstanding.	Merged CRSP - COMPUSTAT
BM	Book to market equity ratio.	Merged CRSP - COMPUSTAT
CAPEXP	Capital expenditures, scaled by total assets.	COMPUSTAT
INTANGIBLE	Intangible assets, scaled by total assets	COMPUSTAT
RND	Research and development expenditures, scaled by total assets.	COMPUSTAT
ROA	Return on assets.	COMPUSTAT
LEV	Total debt to total assets.	COMPUSTAT
PRC	Natural log of the closing share price.	CRSP
TURN	Number of shares traded as a proportion of shares outstanding (in thousands)	CRSP
BIDASK	Closing ask minus the closing bid, divided by the mid-price.	CRSP
STDRET	Natural log of standard deviation of daily stock returns	CRSP
CAR(-50,-21)	Cumulative abnormal return over event dates -50 to -21.	CRSP
CAR(-20,-1)	Cumulative abnormal return over event dates -20 to -1.	CRSP
SP500	Indicator variable coded one if firm is in the Standard and Poor's 500	COMPUSTAT

Appendix B: Variable Definition for Additional Liquidity Tests

In robustness tests, we consider other liquidity measures that are commonly used in the accounting and finance literatures. These include the dollar-weighted average percentage effective spread (*ES*), Kyle's (1985) lambda (*Lambda*), the Amihud (2002) illiquidity measure (*AMI*), quoted spread (*QS*), realized spread (*RS*), price impact (*PI*) and order imbalance (*OI*). Our data source for *ES* and *Lambda* is the daily TAQ database (DTAQ). We use SAS code available on Craig Holden's website (https://kelley.iu.edu/cholden/) and follow the trade-signing approach of Lee and Ready (1991), using contemporaneous quotes to sign trades (e.g., Bessembinder 2003). Calculation of *AMI* employs daily CRSP data.

The effective spread is the difference between an estimate of the true value of the security (the midpoint of the bid and ask) and the actual transaction price, and is computed by comparing the trade price to the prevailing quote midpoint (Bessembinder and Kaufman 1997; Huang and Stoll 1996). For each stock in each day, we calculate the dollar-weighted average percentage effective spread, based on the following definition of the effective spread for a trade at time *t*:

$$Percent\ Effective\ Spread_t = 2 \times D_t \left(Ln\left(P_t\right) - Ln\left(M_t\right) \right)$$

where D_t is an indicator variable that equals +1 if the trade at time *t* is buyer-initiated and −1 if the trade at time *t* is seller-initiated. M_t is the midpoint of the consolidated best bid and offer at the moment of the trade.

Realized spread measures the revenue to liquidity suppliers, including compensation for order processing costs, inventory costs or market power (Brogaard et al. 2015). For each stock in each day, we calculate the percentage realized spread, based on the following definition of the realized spread for a trade at time *t*:

$$RS_t = 2 \times D_t \left(Ln(P_t) - Ln(M_{t+n})\right),$$

where D_t is an indicator variable that equals +1 if the trade at time *t* is buyer-initiated and −1 if the trade at time *t* is seller-initiated. M_{t+5} is the consolidated midpoint of the best bid and offer 5 minutes after the trade.

Price impact reflects the market's assessment of private information conveyed in trades, and is observed by the increase (decrease) in stock price following a customer buy (sell) (Bessembinder and Kaufman 1997). For each stock in each day, we calculate the percentage price impact, based on the following definition of the price impact for a trade at time *t*:

$$PI_t = 2 \times D_t \left(Ln(M_{t+n}) - Ln(M_t)\right),$$

where D_t is an indicator variable that equals +1 if the trade at time t is buyer-initiated and −1 if the trade at time t is seller-initiated. M_{t+5} is the consolidated midpoint of the best bid and offer 5 minutes after the trade and M_t is the midpoint of the consolidated best bid and offer at the moment of the trade.

Order imbalance is measured as | (B – S) / (B +S) | where B is the number of traded shares that are buy market orders and S is the number of traded shares that are sell market orders. Trades are signed following the trade-signing approach of Lee and Ready (1991).

QS is the time-weighted percent quoted spread (during market hours) from DTAQ.

Kyle's (1985) lambda represents the extent to which signed order flow affects a security's price. We follow Hasbrouck (2009) and define *LAMBDA* as the slope of the following regression:

$$r_n = \lambda S_n + \varepsilon_n$$

where r_n is the security's log price change in the n^{th} five-minute period, S_n is the signed square-root of dollar volume in the n^{th} five-minute period, and ε_n is the error term. S_n is defined by

$$S_n = \sum sign(v_{kn} \times SQRT\left(|\, v_{kn} \,|\right)$$

where v_{nk} is the signed dollar volume of the *k*th trade in the *n*th five-minute period.

The Amihud (2002) illiquidity measure, *AMI*, is the ratio of absolute value of daily stock return to the daily dollar trading volume.

Table 1: Descriptive Statistics

Panel A: Dependent variables, control variables and cybersecurity risk mitigation measures

	N	Mean	Median	Std Dev	Min.	Max.
NCyberWords	60,862	1,032	712	1,115	0	15,994
CyberWords	60,862	6.20	6.57	1.76	0.00	9.68
NMitigation	60,862	18.31	11.00	24.96	0.00	483.00
Mitigation	60,862	2.37	2.48	1.18	0.00	6.18
PJR	22,964	0.516	0.515	0.456	-1.005	1.977
ASHVOL	49,997	0.099	0.020	0.657	-1.505	3.048
AOptVol	52,151	-0.047	0.040	1.450	-26.927	6.990
AStkVol	60,846	-0.065	-0.126	0.830	-2.277	3.383
ARange	49,997	0.154	0.015	0.734	-1.569	4.538
AVolatility	49,997	0.425	0.039	1.402	-3.402	16.073
10KWords	60,862	10.824	10.839	0.791	0.000	13.568
10KMitiWords	60,862	5.793	5.782	0.624	1.386	8.669
PREVIOUS	60,862	0.047	0.000	0.211	0.000	1.000
FORECAST	60,862	0.195	0.000	0.396	0.000	1.000
ANALYST	60,860	1.868	1.946	0.874	0.000	3.892
IO	60,862	0.684	0.768	0.287	0.000	1.000
MV	60,849	14.448	14.337	1.674	7.238	20.791
BM	60,849	0.500	0.424	0.353	0.010	2.518
BIDASK	60,849	0.001	0.001	0.002	0.000	0.031
TURN	60,849	8.571	6.576	7.191	0.333	53.645
STDRET	60,849	-4.083	-4.122	0.517	-5.405	-2.423
CAR(-50,-21)	60,849	0.006	0.004	0.121	-1.545	4.487
CAR(-20,1)	60,849	-0.004	-0.002	0.094	-1.019	4.171
PRC	60,849	3.367	3.422	1.013	-2.047	12.682

Panel A reports descriptive statistics of the firms in our sample based upon the test and control variables in our models. The sample period is between 2012 to 2018. All variables are defined in the Appendix B.

Panel B: Pearson correlations

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	***NCyberWords***																							
2	***CyberWords***	0.62																						
3	***NMitigation***	0.89	0.52																					
4	***Mitigation***	0.76	0.83	0.73																				
5	***PJR***	0.06	0.05	0.05	0.06																			
6	***ASHVOL***	0.00	0.00	0.00	0.00	-0.01																		
7	***AOptVol***	-0.08	-0.07	-0.07	-0.08	-0.06	0.06																	
8	***AStkVol***	-0.04	-0.04	-0.03	-0.04	-0.19	0.05	0.20																
9	***ARange***	-0.02	-0.03	-0.01	-0.03	-0.11	0.01	0.15	0.43															
10	***AVolatility***	0.01	0.00	0.01	0.00	-0.01	0.03	0.03	-0.06	0.44														
11	***10KWords***	0.03	0.07	0.02	0.07	-0.06	-0.01	0.03	0.01	0.01	0.01													
12	***10KMitiWords***	0.26	0.26	0.26	0.33	-0.06	-0.01	-0.01	-0.01	-0.01	0.01	0.78												
13	***PREVIOUS***	0.15	0.12	0.12	0.14	0.01	0.01	-0.03	0.02	0.03	0.02	0.04	0.06											
14	***FORECAST***	0.02	0.02	0.02	0.01	0.00	0.00	0.00	-0.04	-0.03	-0.02	0.00	-0.05	0.04										
15	***ANALYST***	0.19	0.18	0.16	0.18	0.05	-0.01	-0.03	0.10	0.07	0.01	0.09	0.09	0.22	0.14									
16	***IO***	-0.02	-0.02	-0.02	-0.04	0.09	0.00	0.07	0.04	0.05	-0.01	0.00	-0.05	0.02	0.11	0.25								
17	***MV***	0.12	0.15	0.09	0.13	0.00	-0.02	-0.06	0.09	0.08	0.03	0.12	0.07	0.30	0.15	0.68	0.20							
18	***BM***	-0.10	-0.07	-0.08	-0.07	-0.06	0.00	-0.01	-0.04	-0.02	-0.03	0.07	0.13	-0.04	-0.10	-0.24	-0.12	-0.27						
19	***BIDASK***	-0.03	-0.05	-0.03	-0.04	-0.08	0.01	0.00	-0.02	-0.04	-0.04	-0.05	0.00	-0.09	-0.13	-0.45	-0.32	-0.57	0.19					
20	***TURN***	0.07	0.06	0.05	0.05	0.00	-0.02	-0.10	-0.19	-0.14	-0.02	0.04	0.03	0.00	0.08	0.28	0.10	0.03	-0.06	-0.16				
21	***STDRET***	0.05	0.02	0.03	0.03	0.00	0.01	-0.09	-0.21	-0.36	-0.11	-0.02	0.03	-0.10	-0.03	-0.13	-0.10	-0.42	0.01	0.27	0.44			
22	***CAR(-50,-21)***	0.00	-0.01	0.00	0.00	0.01	-0.08	0.02	-0.06	0.09	-0.14	0.01	0.01	0.00	0.00	-0.02	0.00	-0.03	0.01	0.04	0.04	0.07		
23	***CAR(-20,-1)***	0.01	0.00	0.01	0.01	0.02	0.04	0.03	0.00	0.07	-0.10	-0.01	-0.01	0.00	0.01	0.00	0.00	0.01	-0.01	0.01	-0.01	0.00	-0.01	
24	***PRC***	0.05	0.09	0.05	0.08	0.06	-0.03	-0.02	0.09	0.08	0.02	0.00	-0.04	0.13	0.10	0.39	0.26	0.69	-0.31	-0.47	-0.05	-0.44	-0.02	0.01

The sample period is fiscal year 2012 to 2018. All correlations with an absolute value larger than 0.008 are significant at the 5% level and all correlations with an absolute value larger than 0.011 are significant at the 1% level. All variables are defined in Appendix B.

Panel C: Alternative cybersecurity risk mitigation measures

	N	Mean	Median	Std Dev	Min.	Max.
IT_Exec	18,529	0.04	0.00	0.28	0.00	11.00
RiskComm	18,529	0.10	0.00	0.30	0.00	1.00
IT_Capexp	18,529	1.68	1.79	0.90	0.00	6.49
NegTone	18,529	0.06	0.06	0.03	0.00	0.33

Panel C reports descriptive statistics of the alternative cybersecurity risk mitigation measures used in our validation tests. The sample period is between 2012 to 2018. All variables are defined in the Appendix B.

Table 2: Logistic Analysis for Data Breach Prediction

	(1) *CyberWords*	(2) *Mitigation*	(3) *NegTone*	(4) *IT_Capexp*	(5) *RiskComm*	(6) *IT_Exec*
Intercept	-18.263	-21.178	-25.364	-16.708	-18.565	-22.563
	(0.736)	(0.708)	(0.678)	(0.998)	(0.245)	(0.686)
CyberMiti	**-0.673**	**-1.395**	**-43.796**	0.207	-0.614	-0.307
	(0.013)	**(0.001)**	**(0.019)**	(0.626)	(0.692)	(0.444)
10KWords	4.572	4.709	4.874	4.437	4.593	4.679
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
10KMitiWords	-2.872	-2.870	-3.199	-2.942	-3.016	-3.079
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
PREVIOUS	-4.255	-4.319	-4.164	-3.826	-4.053	-4.048
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
INDDIR	0.066	0.065	0.028	0.046	0.083	0.071
	(0.711)	(0.719)	(0.880)	(0.790)	(0.643)	(0.691)
BRDSIZE	-6.112	-6.131	-4.872	-5.111	-6.374	-6.230
	(0.065)	(0.066)	(0.151)	(0.107)	(0.053)	(0.059)
IO	1.428	1.557	1.296	1.470	1.378	1.372
	(0.418)	(0.383)	(0.458)	(0.390)	(0.427)	(0.428)
MV	0.774	0.839	0.724	0.662	0.704	0.749
	(0.224)	(0.189)	(0.261)	(0.295)	(0.267)	(0.236)
BM	-1.718	-1.650	-2.258	-1.598	-1.890	-1.788
	(0.189)	(0.218)	(0.098)	(0.208)	(0.138)	(0.159)
CAPEXP	-34.378	-29.466	-27.528	-28.385	-29.795	-30.819
	(0.030)	(0.059)	(0.050)	(0.049)	(0.048)	(0.039)
INTANGIBLE	-19.999	-19.733	-18.415	-18.206	-18.713	-19.621
	(0.011)	(0.014)	(0.018)	(0.017)	(0.016)	(0.013)
RND	-1.768	-3.558	-0.605	-3.448	-2.943	-0.836
	(0.900)	(0.800)	(0.965)	(0.800)	(0.832)	(0.953)
ROA	4.403	4.276	4.380	4.216	4.243	4.359
	(0.248)	(0.256)	(0.260)	(0.266)	(0.274)	(0.260)
LEV	1.458	1.627	2.133	1.176	1.580	1.566
	(0.423)	(0.375)	(0.245)	(0.508)	(0.382)	(0.392)
SP500	-0.449	-0.569	-0.638	0.061	-0.337	-0.231
	(0.665)	(0.593)	(0.533)	(0.954)	(0.740)	(0.819)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	14,305	14,305	14,305	14,305	14,305	14,305
Pseudo-R^2	0.722	0.724	0.731	0.715	0.715	0.715

$$BREACH_{i,t+1} = a + \beta_1 CyberMiti_{i,t} + \beta_2 10KWords_{i,t} + \beta_3 10KMitiWords_{i,t} + \beta_4 PREVIOUS_{i,t} + \beta_5 INDDIR_{i,t} + \beta_6 BRDSIZE_{i,t} + \beta_7 IO_{i,t} + \beta_8 MV_{i,t} + \beta_9 BM_{i,t} + \beta_{10} CAPEXP_{i,t} + \beta_{11} INTANGIBLE_{i,t} + \beta_{12} RND_{i,t} + \beta_{13} ROA_{i,t} + \beta_{14} LEV_{i,t} + \beta_{15} SP500_{i,t} + \Sigma FirmFE + \Sigma YearFE + \varepsilon_{i,t}$$

This table presents the results of cybersecurity risk mitigation measures ($CyberMiti_t$) and the likelihood of a cybersecurity data breach in a firm the following year (t+1). Column (1) presents the results when cybersecurity risk mitigation is captured by the amount of words contained in cybersecurity excerpts disclosed in a firm's 10-K. Column (2) presents the results when cybersecurity risk mitigation is captured by the amount of cybersecurity risk mitigation words or phrases disclosed in a firm's 10-K. Column (3) presents the results when cybersecurity risk mitigation is

captured by the extent of negative tone contained in cybersecurity excerpts disclosed in a firm's 10-K, this is computed as negative words in cybersecurity excerpts disclosed in a firm's 10-K divided by total amount of words contained in cybersecurity excerpts disclosed in a firm's 10-K. Column (4) presents the results when cybersecurity risk mitigation is captured by the natural log plus one of the number of IT software packages words. Column (5) presents the results when cybersecurity risk mitigation is captured by the presence of an IT or risk board committee. Column (6) presents the results when cybersecurity risk mitigation is captured by the number of IT executives in top management of the firm. P-values are presented in parentheses below the coefficients.

Table 3: Firm Response to Peer Cybersecurity Breaches

Panel A: OLS panel regression analysis of peer breaches and one-year change in risk mitigation measures

	(1)	(2)	(3)	(4)	(5)	(6)
	ΔCyberWords	***ΔMitigation***	***ΔNegTone***	***ΔIT_Capexp***	***ΔRiskComm***	***ΔIT_Exec***
Intercept	-136.283	2.017	0.016	1.275	-0.196	-0.042
	(-0.86)	(0.70)	(1.65)	(0.96)	(-1.30)	(-0.75)
PEER	**23.241**	0.335	0.000	0.171	0.004	-0.005
	(2.06)	(1.47)	(-0.28)	(1.17)	(0.47)	(-0.76)
10KWords	-0.098	-0.534	0.002	0.05	-0.033	0.003
	(0.00)	(-1.20)	(1.05)	(0.25)	(-1.44)	(0.39)
10KMitiWords	31.968	0.801	-0.005	-0.631	0.076	0.001
	(1.69)	(1.81)	(-2.94)	(-3.01)	(3.45)	(0.10)
PREVIOUS	93.957	1.029	-0.001	-0.333	0.058	0.003
	(1.34)	(0.92)	(-0.96)	(-0.74)	(1.32)	(0.14)
INDDIR	-0.351	0.020	0.000	-0.013	0.004	0.000
	(-0.26)	(0.71)	(-0.90)	(-1.13)	(2.70)	(-0.09)
BRDSIZE	-13.522	-0.266	0.001	0.156	-0.009	0.002
	(-0.79)	(-0.73)	(0.67)	(1.02)	(-0.45)	(0.32)
IO	3.630	0.194	0.001	0.017	0.055	0.004
	(0.28)	(0.88)	(0.74)	(0.18)	(3.41)	(0.97)
MV	14.521	0.326	0.000	0.031	0.028	0.001
	(3.97)	(4.94)	(-0.66)	(0.97)	(5.88)	(0.88)
BM	1.053	0.053	0.000	9E-04	0.015	-0.001
	(0.32)	(0.85)	(0.21)	(0.03)	(1.74)	(-0.97)
CAPEXP	214.985	4.367	0.010	1.097	-0.112	-0.003
	(2.30)	(2.55)	(1.23)	(1.27)	(-1.04)	(-0.09)
RND	86.443	1.967	0.002	0.089	-0.023	-0.010
	(4.05)	(5.43)	(0.83)	(0.45)	(-0.65)	(-1.20)
ROA	166.092	2.843	0.013	0.147	-0.193	-0.006
	(2.33)	(2.02)	(1.37)	(0.20)	(-3.74)	(-0.38)
LEV	-21.272	-0.770	0.003	-0.128	-0.149	0.006
	(-0.63)	(-1.24)	(1.11)	(-0.43)	(-5.78)	(0.65)
INTANGIBLE	6.311	0.274	0.001	-0.156	-0.012	-0.006
	(0.42)	(0.88)	(0.91)	(-1.21)	(-0.6)	(-1.14)
SP500	-9.392	-0.339	-0.001	-0.315	-0.044	0.005
	(-0.71)	(-1.33)	(-1.35)	(-2.17)	(-2.33)	(0.88)
FF48 FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	14,305	14,305	14,305	14,305	14,305	14,305
Adj-R^2	0.027	0.039	0.018	0.013	0.269	0.004

Panel B: OLS panel regression analysis of peer breaches and two-year change in risk mitigation measures

	(1)	(2)	(3)	(4)	(5)	(6)
	ΔCyberWords	***ΔMitigation***	***ΔNegTone***	***ΔIT_Capexp***	***ΔRiskComm***	***ΔIT_Exec***
Intercept	-380.046	-0.446	0.088	3.237	-0.338	-0.119
	(-1.40)	(-0.08)	(0.89)	(1.20)	(-1.98)	(-1.35)
PEER	**48.213**	**0.900**	-0.005	0.238	0.011	-0.003
	(2.46)	**(2.44)**	(-1.30)	(1.30)	(1.08)	(-0.39)
10KWords	22.309	-0.637	0.018	0.08	-0.027	0.007
	(0.61)	(-0.79)	(1.55)	(0.22)	(-1.06)	(0.63)
10KMitiWords	46.382	1.730	-0.042	-1.076	0.085	-0.008
	(1.35)	(2.16)	(-3.78)	(-2.82)	(3.57)	(-0.85)
PREVIOUS	228.021	3.084	0.006	-0.283	0.055	-0.012
	(1.38)	(1.14)	(0.66)	(-0.37)	(1.19)	(-0.38)
INDDIR	-0.825	0.027	-0.000	-0.032	0.004	-0.001
	(-0.29)	(0.42)	(-1.14)	(-1.22)	(2.26)	(-0.63)
BRDSIZE	-22.084	-0.480	0.013	0.454	-0.001	0.016
	(-0.61)	(-0.60)	(0.98)	(1.04)	(-0.03)	(1.17)
IO	23.145	0.578	-0.012	-0.028	0.060	0.014
	(0.97)	(1.29)	(-1.47)	(-0.14)	(3.39)	(1.72)
MV	16.396	0.387	-0.009	0.019	0.028	0.001
	(2.38)	(3.11)	(-4.03)	(0.33)	(5.3)	(0.51)
BM	-0.982	0.087	0.003	-5E-04	0.013	-0.002
	(-0.21)	(0.91)	(1.72)	(-0.01)	(1.55)	(-1.13)
CAPEXP	344.516	7.737	-0.072	2.071	-0.116	-0.061
	(2.44)	(2.80)	(-0.97)	(0.97)	(-1.02)	(-1.08)
INTANGIBLE	133.305	3.135	-0.000	0.189	-0.023	-0.022
	(3.07)	(4.36)	(-0.01)	(0.40)	(-0.58)	(-1.42)
RND	316.110	6.780	0.058	-2.41	-0.220	0.039
	(2.52)	(2.66)	(0.91)	(-1.33)	(-3.52)	(1.00)
ROA	-63.159	-0.725	-0.029	-1.379	-0.148	0.030
	(-1.18)	(-0.66)	(-0.95)	(-1.75)	(-4.95)	(1.6)
LEV	-3.349	0.111	0.022	-0.174	-0.006	-0.006
	(-0.12)	(0.19)	(1.97)	(-0.62)	(-0.27)	(-0.66)
SP500	13.345	0.038	0.014	-0.439	-0.046	0.020
	(0.49)	(0.07)	(1.94)	(-1.7)	(-2.25)	(1.76)
FF48 FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	11,011	11,011	11,011	11,011	11,011	11,011
Adj-R²	0.046	0.066	0.034	0.017	0.284	0.008

$$\Delta CyberMiti_{i,t} = a + \beta_1 PEER_{i,t} + \beta_2 10KWords_{i,t} + \beta_3 10KMitiWords_{i,t} + \beta_4 PREVIOUS_{i,t} + \beta_5 INDDIR_{i,t} + \beta_6 BRDSIZE_{i,t} + \beta_7 IO_{i,t} + \beta_8 MV_{i,t} + \beta_9 BM_{i,t} + \beta_{10} CAPEXP_{i,t} + \beta_{11} INTANGIBLE_{i,t} + \beta_{12} RND_{i,t} + \beta_{13} ROA_{i,t} + \beta_{14} LEV_{i,t} + \beta_{15} SP500_{i,t} + \Sigma FF48FE + \Sigma YearFE + \varepsilon_{i,t}$$

Panel A (B) presents the results of a one(two)-year change in cybersecurity risk mitigation measures in peer firms after a breach in a focal firm. The dependent variable is the one(two)-year change in cybersecurity risk mitigation measures. Peer firms are classified based on TNIC. Column (1) presents the results when cybersecurity risk mitigation is captured by the amount of words contained in cybersecurity excerpts disclosed in a firm's 10-K. Column (2) presents the results when cybersecurity risk mitigation is captured by the amount of cybersecurity risk mitigation words or phrases disclosed in a firm's 10-K. Column (3) presents the results when cybersecurity risk mitigation is captured by the extent of negative tone contained in cybersecurity excerpts disclosed in a firm's 10-K, this is computed as negative

words in cybersecurity excerpts disclosed in a firm's 10-K divided by total amount of words contained in cybersecurity excerpts disclosed in a firm's 10-K. Column (4) presents the results when cybersecurity risk mitigation is captured by the natural log plus one of the number of IT software packages words. Column (5) presents the results when cybersecurity risk mitigation is captured by the presence of an IT or risk board committee. Column (6) presents the results when cybersecurity risk mitigation is captured by the number of IT executives in top management of the firm. Standard errors are clustered by firm. T-statistics are presented in parentheses below the coefficients.

Table 4: OLS Panel Regression Analysis of Cybersecurity Risk Mitigation and the Price Jump Ratio

	(1)		(2)		(3)		(4)		(5)		(6)	
	PJR		*PJR based on (0,2)/(-10,2)*		*PJR*		*PJR*		*PJR*		*PJR_Volume*	
	Weller criterion		Weller criterion		Excl. bottom 50%		Excl. bottom 25%		Incl. All		Incl. All	
	CyberWords	*Mitigation*	*CyberWords*	*Mitigation*	*CyberWords*	*Mitigation*	*CyberWords*	*Mitigation*	*CyberWords*	*Mitigation*	*CyberWords*	*Mitigation*
Intercept	0.014	0.009	0.006	0.002	0.012	0.008	0.017	0.015	0.003	0.007	0.081	0.075
	(3.69)	(2.28)	(1.69)	(0.58)	(4.15)	(2.24)	(4.94)	(3.09)	(0.23)	(0.49)	(11.14)	(9.04)
CyberMiti	**0.008**	**0.021**	**0.008**	**0.019**	**0.009**	**0.020**	**0.007**	0.011	0.012	0.005	**0.018**	**0.035**
	(2.03)	**(4.39)**	**(2.36)**	**(3.71)**	**(2.99)**	**(3.92)**	**(2.15)**	(1.35)	(1.10)	(0.18)	**(3.83)**	**(4.12)**
10KWords	-0.003	0.003	0.023	0.027	-0.037	-0.032	-0.004	-0.002	0.182	0.180	-0.058	-0.050
	(-0.16)	(0.15)	(0.82)	(0.98)	(-1.35)	(-1.20)	(-0.1)	(-0.05)	(2.29)	(2.24)	(-1.96)	(-1.74)
10KMitiWords	0.009	0.002	-0.021	-0.027	0.048	0.042	-0.001	-0.004	-0.142	-0.139	0.088	0.079
	(0.51)	(0.08)	(-0.73)	(-0.94)	(1.69)	(1.50)	(-0.03)	(-0.1)	(-1.63)	(-1.56)	(2.97)	(2.68)
PREVIOUS	0.045	0.044	-0.026	-0.028	0.018	0.016	0.016	0.015	0.344	0.342	0.095	0.092
	(0.83)	(0.8)	(-0.57)	(-0.59)	(0.33)	(0.29)	(0.24)	(0.22)	(2.06)	(2.04)	(2.00)	(1.93)
FORECAST	-0.050	-0.050	-0.034	-0.034	-0.058	-0.058	-0.054	-0.054	-0.032	-0.032	-0.137	-0.137
	(-4.42)	(-4.4)	(-3.62)	(-3.61)	(-4.63)	(-4.64)	(-4.95)	(-4.94)	(-0.73)	(-0.73)	(-9.42)	(-9.4)
ANALYST	0.037	0.037	0.040	0.041	0.045	0.045	0.044	0.044	0.011	0.011	0.096	0.097
	(3.98)	(4.04)	(4.41)	(4.44)	(3.67)	(3.67)	(3.1)	(3.11)	(0.27)	(0.27)	(7.17)	(7.16)
IO	0.109	0.109	0.123	0.123	0.102	0.102	0.127	0.127	0.124	0.122	0.179	0.179
	(4.95)	(4.97)	(6.1)	(6.08)	(4.36)	(4.35)	(5.44)	(5.44)	(1.76)	(1.72)	(3.25)	(3.24)
MV	-0.002	-0.003	-0.007	-0.007	-0.018	-0.019	-0.023	-0.023	-0.020	-0.019	-0.011	-0.012
	(-0.33)	(-0.46)	(-0.66)	(-0.73)	(-1.87)	(-1.92)	(-1.73)	(-1.73)	(-0.68)	(-0.66)	(-0.71)	(-0.76)
BIDASK	-0.063	-0.062	-0.080	-0.079	-0.122	-0.121	-0.115	-0.114	-0.191	-0.190	-0.110	-0.109
	(-2.44)	(-2.4)	(-2.75)	(-2.72)	(-4.04)	(-3.99)	(-3.12)	(-3.1)	(-2.03)	(-2.02)	(-2.13)	(-2.10)
STDRET	0.006	0.006	0.014	0.014	-0.020	-0.019	-0.004	-0.004	-0.023	-0.022	-0.057	-0.057
	(0.56)	(0.52)	(1.25)	(1.23)	(-2.11)	(-2.10)	(-0.32)	(-0.31)	(-1.03)	(-0.99)	(-3.07)	(-3.07)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Qtr FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	22,964	22,964	29,172	29,172	30,240	30,240	45,365	45,365	60,473	60,473	60,473	60,473
Adj R2	0.013	0.013	0.007	0.007	0.009	0.009	0.005	0.005	0.001	0.001	0.02	0.021

$$PJR_{i,t} = a + \beta_1 CyberMiti_{i,t} + \beta_2 10KWords_{i,t} + \beta_3 10KMitiWords_{i,t} + \beta_4 PREVIOUS_{i,t} + \beta_5 FORECAST_{i,t} + \beta_6 ANALYST_{i,t} + \beta_7 IO_{i,t} + \beta_8 MV_{i,t} + \beta_9 BIDASK_{i,t} + \beta_{10} STDRET_{i,t} + \Sigma YearQrtFE + \Sigma FirmFE + \varepsilon_{i,t}$$

This table presents the results of cybersecurity risk mitigation and the price jump ratio (*PJR*). *PJR* is calculated as CAR_{it} (0,2) / CAR_{it} (-21,+2). Column (1) reports the results using *PJR*. We follow the criterion in Weller (2018) and retain observations where the absolute value of *CAR*(-21,+2) is larger than the daily return volatility in the preceding month multiplied with the square root of 24. Column (2) reports the results where the price jump ratio is calculated as CAR_{it} (0,2) / CAR_{it} (-10,+2), and retains observations following Weller's (2018) criterion. Column (3) presents results of *PJR* when we drop observations in the lower median of absolute *CAR*(-21,+2). Column (4) presents results of *PJR* when we drop observations with an absolute *CAR*(-21,+2) below the 25th percentile. Column (5) presents results of *PJR* using the full sample. Column (6) reports the results using a volume based PJR where PJR_Volume_{it} is $TradingVolume_{it}$ (0,+2) / $TradingVolume_{it}$ (-21,+2). Standard errors are clustered by firm and quarter-year. T-statistics are presented in parentheses below the coefficients.

Table 5: Cybersecurity Risk Mitigation and Informativeness of Short Selling

Panel A: OLS Panel regression analysis of cybersecurity risk mitigation and short sales leading up to earnings announcements (20-day window)

	(1)		(2)		(3)		(4)	
	CyberWords	***Mitigation***	***CyberWords***	***Mitigation***	***CyberWords***	***Mitigation***	***CyberWords***	***Mitigation***
Intercept	0.002	-0.008	-0.002	-0.018	0.025	0.012	0.014	0.004
	(0.03)	(-0.12)	(-0.03)	(-0.27)	(0.39)	(0.19)	(0.23)	(0.06)
CyberMiti	0.013	0.034	0.000	0.033	0.004	0.031	0.007	0.030
	(0.71)	(1.06)	(-0.01)	(1.02)	(0.23)	(0.96)	(0.38)	(0.92)
***ASHVOL*(-20,-1)**	-0.270	-0.151	-0.265	-0.141	-0.320	-0.157	-0.229	-0.122
	(-3.12)	(-3.11)	(-3.09)	(-2.89)	(-4.10)	(-3.34)	(-2.69)	(-2.54)
ASHVOL× CyberMiti	**0.036**	**0.044**	**0.035**	**0.041**	**0.042**	**0.044**	**0.031**	**0.037**
	(2.74)	**(2.53)**	**(2.75)**	**(2.38)**	**(3.62)**	**(2.59)**	**(2.41)**	**(2.14)**
10KWords	-0.021	-0.018	-0.035	-0.031	-0.003	0.001	-0.012	-0.009
	(-0.50)	(-0.43)	(-0.82)	(-0.72)	(-0.06)	(0.02)	(-0.28)	(-0.20)
10KMitiWords	0.090	0.084	0.104	0.095	0.035	0.028	0.089	0.082
	(1.26)	(1.17)	(1.44)	(1.31)	(0.47)	(0.37)	(1.23)	(1.13)
PREVIOUS	0.156	0.151	0.133	0.130	0.283	0.279	0.174	0.171
	(0.75)	(0.72)	(0.71)	(0.69)	(1.26)	(1.24)	(0.83)	(0.81)
FORECAST	0.042	0.042	0.006	0.006	-0.021	-0.021	0.062	0.062
	(1.12)	(1.12)	(0.17)	(0.16)	(-0.54)	(-0.54)	(1.66)	(1.66)
MV	-0.422	-0.424	-0.351	-0.358	-0.485	-0.489	-0.380	-0.384
	(-4.09)	(-4.11)	(-3.35)	(-3.41)	(-4.80)	(-4.84)	(-3.69)	(-3.72)
BM	0.228	0.228	0.183	0.180	0.358	0.357	0.262	0.261
	(1.95)	(1.95)	(1.51)	(1.49)	(3.34)	(3.33)	(2.27)	(2.26)
PRC	0.216	0.217	-0.119	-0.114	0.583	0.587	0.168	0.170
	(2.07)	(2.08)	(-1.12)	(-1.07)	(5.64)	(5.68)	(1.60)	(1.62)
BIDASK	0.110	0.112	0.061	0.063	0.144	0.145	0.118	0.119
	(0.86)	(0.87)	(0.45)	(0.46)	(1.12)	(1.13)	(0.90)	(0.91)
TURN	-0.042	-0.043	-0.026	-0.027	-0.095	-0.096	-0.032	-0.033
	(-1.10)	(-1.12)	(-0.63)	(-0.64)	(-2.67)	(-2.69)	(-0.86)	(-0.88)
STDRET	-0.755	-0.709	-1.319	-1.325	-0.954	-0.936	-0.702	-0.680
	(-0.36)	(-0.34)	(-0.59)	(-0.59)	(-0.49)	(-0.48)	(-0.34)	(-0.33)
***CAR*(-50,-21)**	20.138	20.167	16.333	16.379	17.591	17.637	18.613	18.645
	(5.83)	(5.83)	(4.40)	(4.41)	(5.64)	(5.66)	(5.40)	(5.41)
***CAR*(-20,-1)**	37.405	37.350	33.809	33.763	40.210	40.141	35.195	35.150
	(10.16)	(10.15)	(8.65)	(8.63)	(11.74)	(11.72)	(9.62)	(9.60)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Qtr FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	49,432	49,432	49,432	49,432	44,996	44,996	49,432	49,432
Adj R2	0.007	0.007	0.009	0.009	0.007	0.007	0.006	0.006

Panel B: OLS panel regression analysis of cybersecurity risk mitigation and short selling leading up to earnings announcements (one-week windows)

	(1)		(2)		(3)		(4)		(5)		(6)	
	Week = -1: Day (-5,-1)		*Week = -2: Day (-10,-6)*		*Week = -3: Day (-15,-11)*		*Week = -4: Day (-20,-16)*		*Week = -5: Day (-25,-21)*		*Week = -6: Day (-26,-30)*	
	CyberWords	***Mitigation***	***CyberWords***	***Mitigation***	***CyberWords***	***Mitigation***	***CyberWords***	***Mitigation***	***CyberWords***	***Mitigation***	***CyberWords***	***Mitigation***
Intercept	0.014	0.004	0.017	0.007	0.024	0.014	0.032	0.022	0.038	0.028	0.027	0.016
	(0.22)	(0.06)	(0.27)	(0.11)	(0.37)	(0.22)	(0.50)	(0.33)	(0.60)	(0.43)	(0.42)	(0.26)
CyberMiti	0.013	0.035	0.014	0.037	0.013	0.034	0.012	0.036	0.013	0.037	0.014	0.038
	(0.72)	(1.12)	(0.80)	(1.15)	(0.73)	(1.08)	(0.68)	(1.11)	(0.76)	(1.15)	(0.81)	(1.19)
ASHVOL(Week)	-0.119	-0.073	-0.176	-0.094	-0.143	-0.075	-0.142	-0.088	-0.088	-0.055	0.040	0.024
	(-1.89)	(-2.16)	(-2.87)	(-2.71)	(-2.27)	(-2.05)	(-2.18)	(-2.38)	(-1.33)	(-1.46)	(0.59)	(0.59)
ASHVOL×CyberMiti	0.015	0.019	**0.023**	**0.025**	**0.022**	**0.028**	0.019	**0.028**	0.009	0.010	-0.005	-0.005
	(1.54)	(1.57)	**(2.44)**	**(2.00)**	**(2.25)**	**(2.12)**	(1.93)	**(2.03)**	(0.88)	(0.69)	(-0.43)	(-0.32)
10KWords	-0.022	-0.019	-0.023	-0.020	-0.021	-0.018	-0.022	-0.019	-0.022	-0.018	-0.022	-0.019
	(-0.52)	(-0.44)	(-0.54)	(-0.47)	(-0.49)	(-0.42)	(-0.52)	(-0.45)	(-0.51)	(-0.44)	(-0.53)	(-0.45)
10KMitiWords	0.094	0.087	0.094	0.087	0.090	0.084	0.093	0.086	0.092	0.085	0.093	0.087
	(1.29)	(1.20)	(1.30)	(1.21)	(1.24)	(1.15)	(1.28)	(1.18)	(1.27)	(1.17)	(1.29)	(1.19)
PREVIOUS	0.149	0.145	0.151	0.148	0.142	0.140	0.163	0.160	0.149	0.147	0.148	0.146
	(0.70)	(0.68)	(0.71)	(0.69)	(0.67)	(0.65)	(0.77)	(0.75)	(0.70)	(0.69)	(0.69)	(0.68)
FORECAST	0.039	0.039	0.042	0.042	0.042	0.042	0.041	0.041	0.040	0.039	0.041	0.041
	(1.03)	(1.04)	(1.12)	(1.11)	(1.11)	(1.11)	(1.09)	(1.08)	(1.06)	(1.05)	(1.10)	(1.09)
MV	-0.447	-0.450	-0.441	-0.443	-0.435	-0.437	-0.437	-0.441	-0.428	-0.431	-0.433	-0.436
	(-4.33)	(-4.36)	(-4.25)	(-4.28)	(-4.21)	(-4.23)	(-4.23)	(-4.27)	(-4.13)	(-4.16)	(-4.19)	(-4.22)
BM	0.241	0.241	0.234	0.235	0.230	0.230	0.233	0.232	0.239	0.239	0.234	0.234
	(2.06)	(2.06)	(2.01)	(2.01)	(1.97)	(1.97)	(1.99)	(1.99)	(2.05)	(2.05)	(2.00)	(2.00)
PRC	0.225	0.226	0.211	0.212	0.206	0.208	0.210	0.212	0.209	0.210	0.201	0.202
	(2.15)	(2.17)	(2.01)	(2.03)	(1.97)	(1.98)	(2.00)	(2.03)	(1.99)	(2.01)	(1.92)	(1.94)
BIDASK	0.113	0.114	0.112	0.114	0.112	0.113	0.110	0.111	0.114	0.115	0.126	0.127
	(0.89)	(0.89)	(0.87)	(0.89)	(0.87)	(0.88)	(0.85)	(0.86)	(0.89)	(0.90)	(0.98)	(0.99)
TURN	-0.046	-0.047	-0.045	-0.045	-0.044	-0.045	-0.044	-0.044	-0.047	-0.047	-0.038	-0.038
	(-1.21)	(-1.23)	(-1.17)	(-1.18)	(-1.16)	(-1.17)	(-1.16)	(-1.16)	(-1.23)	(-1.23)	(-1.01)	(-1.01)
STDRET	-0.104	-0.084	-0.395	-0.378	-0.101	-0.066	-0.253	-0.227	-0.142	-0.138	-0.220	-0.222
	(-0.05)	(-0.04)	(-0.19)	(-0.18)	(-0.05)	(-0.03)	(-0.12)	(-0.11)	(-0.07)	(-0.07)	(-0.10)	(-0.10)
CAR(-50,Week-1)	20.143	20.160	19.632	19.641	19.849	19.870	20.255	20.242	20.394	20.386	16.189	16.186
	(5.83)	(5.84)	(5.68)	(5.68)	(5.74)	(5.74)	(5.86)	(5.85)	(5.90)	(5.90)	(5.52)	(5.52)
CAR (Week)	11.655	11.648	8.711	8.666	7.908	7.904	8.627	8.634	8.850	8.847	4.413	4.403
	(6.49)	(6.49)	(4.56)	(4.53)	(4.50)	(4.49)	(4.87)	(4.88)	(5.09)	(5.09)	(2.55)	(2.54)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Qtr FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	49,432	49,432	49,432	49,432	49,432	49,432	49,432	49,432	49,432	49,432	49,432	49,432
Adj R2	0.005	0.005	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004

$$RSURPRISE_{i,t} = \alpha + \beta_1 CyberMiti_{i,t} + \beta_2 ASHVOL(-20,-1)_{i,t} + \beta_3 CyberMiti_{i,t} \times ASHVOL(-20,-1)_{i,t} + \beta_4 10KWords_{i,t} + \beta_5 10KMitiWords_{i,t} + \beta_6 PREVIOUS_{i,t} + \beta_7 FORECAST_{i,t} + \beta_8 MV_{i,t} +$$
$$\beta_9 BM_{i,t} + \beta_{10} PRC_{i,t} + \beta_{11} BIDASK_{i,t} + \beta_{12} TURN_{i,t} + \beta_{13} STDRET_{i,t} + \beta_{14} CAR(-50,-21)_{i,t} + \beta_{15} CAR\ (-20,-1)_{i,t} + \Sigma FirmFE + \Sigma YearQrt + \varepsilon_{i,t}$$

This table presents the results of cybersecurity risk mitigation and the informativeness of short selling. Panel A uses a 20-day pre-earnings announcement window to measure *ASHVOL(-20,-1)*. Column (1) presents the base model where earnings surprise is calculated as the difference between the actual earnings per share and the most recent average earnings per share forecast from I/B/E/S, scaled by the absolute value of this most recent average earnings per share forecast. Column (2) is the quarterly rank decile of the earnings surprise, with earnings surprise defined as the difference between the actual earnings per share and the most recent average earnings per share forecast from I/B/E/S, scaled by the stock price 10 days before the earnings announcement. Column (3) reports the results when the earnings surprise measure is calculated as the difference between the actual earnings per share and the most recent average earnings per share estimate across analysts, scaled by the standard deviation across analyst estimates. Column (4) reports the results when the earnings surprise measure is calculated as the difference between the actual earnings per share and the most recent median earnings per share estimate across analysts, scaled by the absolute value of this most recent median earnings per share forecast. Panel B presents our results using various one-week windows [weeks(-1,-6)] leading up to earnings announcement. Earnings surprise in panel B is calculated as the difference between the actual earnings per share and the most recent average earnings per share forecast scaled by the absolute value of this most recent average earnings per share forecast. Standard errors are clustered by firm and quarter-year. T-statistics are presented in parentheses below the coefficients.

Table 6: Cybersecurity Risk Mitigation and Informed Trading Proxies in Pre-Earnings Announcements
Panel A: OLS panel regression analysis of cybersecurity risk mitigation and informed trading proxies leading up to earnings announcements (20-day window)

	(1)		(2)		(3)		(4)	
	AOptVol		*AStkVol*		*ARange*		*AVolatility*	
	CyberWords	***Mitigation***	***CyberWords***	***Mitigation***	***CyberWords***	***Mitigation***	***CyberWords***	***Mitigation***
Intercept	-3.829	-3.811	-0.130	-0.126	-0.264	-0.261	-0.372	-0.364
	(-39.37)	(-39.28)	(-8.02)	(-7.72)	(-22.31)	(-21.83)	(-20.74)	(-20.16)
CyberMiti	**-0.009**	**-0.045**	**-0.016**	**-0.028**	**-0.010**	**-0.018**	**-0.010**	**-0.028**
	(-2.05)	**(-4.05)**	**(-4.94)**	**(-4.37)**	**(-3.05)**	**(-3.18)**	**(-1.97)**	**(-2.95)**
10KWords	0.087	0.082	0.012	0.011	-0.006	-0.007	-0.014	-0.016
	(2.13)	(1.99)	(1.45)	(1.26)	(-0.72)	(-0.87)	(-1.08)	(-1.27)
10KMitiWords	-0.132	-0.122	-0.023	-0.020	0.007	0.009	0.021	0.026
	(-2.78)	(-2.54)	(-1.60)	(-1.37)	(0.55)	(0.74)	(0.98)	(1.22)
PREVIOUS	-0.307	-0.305	0.006	0.009	0.014	0.016	0.058	0.059
	(-3.40)	(-3.37)	(0.13)	(0.20)	(0.40)	(0.44)	(0.94)	(0.95)
FORECAST	-0.067	-0.066	-0.033	-0.033	-0.032	-0.032	-0.012	-0.012
	(-4.60)	(-4.59)	(-3.19)	(-3.20)	(-3.97)	(-3.98)	(-0.89)	(-0.88)
PRC	-0.001	-0.001	0.000	0.000	0.000	0.000	0.000	0.000
	(-3.16)	(-3.16)	(4.81)	(4.67)	(1.33)	(1.33)	(1.64)	(1.64)
MV	-0.059	-0.057	0.020	0.021	-0.112	-0.111	-0.017	-0.016
	(-2.45)	(-2.39)	(1.78)	(1.84)	(-8.51)	(-8.45)	(-1.10)	(-1.02)
TURN	-0.049	-0.049	-0.085	-0.085	-0.002	-0.002	0.020	0.020
	(-1.76)	(-1.76)	(-2.40)	(-2.40)	(-0.22)	(-0.22)	(1.68)	(1.68)
STDRET	-4.822	-4.815	-4.359	-4.372	-10.241	-10.247	-6.550	-6.551
	(-3.80)	(-3.81)	(-3.06)	(-3.07)	(-6.09)	(-6.10)	(-6.37)	(-6.37)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Qtr FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	52,120	52,120	60,792	60,792	49,996	49,996	49,996	49,996
Adj R2	0.354	0.355	0.139	0.139	0.268	0.268	0.161	0.161

Panel B: Coefficients of *CyberMiti* from OLS panel regression analyses of cybersecurity risk mitigation and informed trading proxies leading up to earnings announcements (various windows)

	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)	
	Week -4, -1		Week -1		Week -2		Week -3		Week -4		Week -5		Week -6		Week +1	
	CyberWords	*Mitigation*	*CyberWords*	*Mitigation*	*CyberWords*	*Mitigation*	*CyberWords*	*Mitigation*	*CyberWords*	*Mitigation*	*CyberWords*	*Mitigation*	*CyberWords*	*Mitigation*	*CyberWords*	*Mitigation*
AOptVol	**-0.009**	**-0.045**	-0.004	**-0.036**	**-0.012**	**-0.052**	-0.010	**-0.044**	**-0.011**	**-0.050**	**-0.012**	**-0.048**	**-0.010**	**-0.020**	-0.001	**-0.030**
	(-2.05)	**(-4.05)**	(-0.72)	**(-2.66)**	**(-2.33)**	**(-4.01)**	(-1.87)	**(-3.42)**	**(-2.17)**	**(-4.62)**	**(-2.69)**	**(-5.14)**	**(-2.88)**	**(-2.76)**	(-0.18)	**(-2.14)**
AStkVol	**-0.016**	**-0.028**	-0.007	-0.012	**-0.017**	**-0.032**	**-0.023**	**-0.042**	**-0.017**	**-0.027**	**-0.013**	**-0.019**	-0.002	-0.002	0.001	0.005
	(-4.94)	**(-4.37)**	(-1.78)	(-1.44)	**(-4.16)**	**(-4.00)**	**(-5.75)**	**(-5.29)**	**(-4.20)**	**(-3.33)**	**(-3.15)**	**(-2.43)**	(-0.60)	(-0.26)	(0.21)	(0.52)
ARange	**-0.010**	**-0.018**	-0.007	-0.009	**-0.009**	**-0.020**	**-0.012**	**-0.023**	**-0.012**	**-0.021**	-0.006	-0.014	-0.002	-0.007	0.002	0.017
	(-3.05)	**(-3.18)**	(-1.52)	(-1.08)	**(-2.04)**	**(-2.54)**	**(-2.78)**	**(-2.89)**	**(-3.01)**	**(-2.90)**	(-1.37)	(-1.76)	(-0.56)	(-1.12)	(0.36)	(1.77)
AVol	**-0.010**	**-0.028**	-0.008	**-0.030**	-0.008	**-0.030**	-0.009	-0.014	**-0.017**	**-0.040**	-0.004	-0.015	0.001	0.001	0.004	0.016
	(-1.96)	**(-2.95)**	(-1.06)	**(-2.16)**	(-1.24)	**(-2.45)**	(-1.47)	(-1.24)	**(-2.92)**	**(-3.96)**	(-0.86)	(-1.64)	(0.32)	(0.10)	(0.51)	(1.19)
AMI	**0.004**	0.002	0.004	0.003	0.005	-0.008	0.004	0.007	0.005	0.007	0.005	**0.012**	0.002	0.000	0.004	**0.010**
	(1.99)	(0.49)	(1.39)	(0.58)	(1.36)	(-1.39)	(1.37)	(1.13)	(1.63)	(1.25)	(1.77)	**(2.31)**	(0.67)	(0.04)	(1.43)	**(1.97)**
ES	**-0.008**	**-0.016**	-0.006	**-0.019**	-0.008	-0.016	**-0.013**	-0.016	-0.004	**-0.015**	-0.008	-0.016	0.003	0.008	0.002	-0.003
	(-2.14)	**(-2.58)**	(-1.21)	**(-2.09)**	(-1.49)	(-1.81)	**(-2.95)**	(-1.95)	(-1.02)	**(-1.97)**	(-1.76)	(-1.87)	(0.58)	(1.01)	(0.36)	(-0.33)
QS	**-0.019**	**-0.037**	-0.009	-0.010	**-0.020**	**-0.045**	**-0.026**	**-0.053**	**-0.023**	**-0.051**	-0.009	-0.013	-0.001	-0.003	-0.001	0.014
	(-2.88)	**(-3.09)**	(-1.01)	(-0.65)	**(-2.44)**	**(-3.12)**	**(-3.59)**	**(-3.89)**	**(-3.18)**	**(-3.79)**	(-1.41)	(-1.15)	(-0.20)	(-0.30)	(-0.07)	(0.81)
RS	0.001	0.003	0.002	0.000	0.000	0.002	0.005	0.008	-0.001	0.002	-0.004	0.004	0.003	0.009	-0.003	-0.007
	(0.63)	(0.75)	(0.68)	(-0.04)	(-0.10)	(0.36)	(1.51)	(1.43)	(-0.32)	(0.35)	(-1.19)	(0.64)	(1.11)	(1.69)	(-1.01)	(-1.32)
PI	**-0.005**	**-0.011**	-0.005	-0.011	-0.003	-0.010	**-0.011**	**-0.016**	-0.002	-0.008	-0.003	**-0.016**	-0.001	-0.004	0.003	0.004
	(-2.24)	**(-2.80)**	(-1.50)	(-1.86)	(-0.75)	(-1.77)	**(-3.28)**	**(-2.75)**	(-0.57)	(-1.50)	(-0.91)	**(-2.58)**	(-0.16)	(-0.56)	(1.01)	(0.59)
OI	-0.002	-0.003	0.001	-0.001	-0.003	-0.004	0.000	-0.005	-0.004	-0.003	-0.002	-0.003	0.001	0.000	-0.001	0.002
	(-0.78)	(-0.77)	(0.16)	(-0.11)	(-0.96)	(-0.64)	(-0.08)	(-0.88)	(-1.25)	(-0.48)	(-0.63)	(-0.47)	(0.19)	(-0.05)	(-0.39)	(0.35)
Lambda	**-0.010**	**-0.015**	-0.008	-0.009	-0.007	-0.008	-0.009	-0.014	-0.008	-0.016	-0.006	-0.018	-0.007	-0.004	0.007	0.011
	(-2.67)	**(-2.57)**	(-1.66)	(-1.18)	(-1.40)	(-0.92)	(-1.87)	(-1.74)	(-1.62)	(-1.95)	(-1.47)	(-2.31)	(-1.60)	(-0.56)	(1.56)	(1.46)

$$A\_ASINFO_{1\text{–}4,i,q,t} = a + \beta_1 CyberMiti_{i,q} + \beta_2 10KWords_{i,q} + \beta_2 10KMitiWords_{i,q} + \beta_3 PREVIOUS_{i,q} + \beta_4 FORECAST_{i,q} + \beta_5 PRC_{i,q} + \beta_6 MV_{i,q} + \beta_7 TURN_{i,q} + \beta_8 STDRET_{i,q} + \Sigma FirmFE + \Sigma YearQrtFE + \varepsilon_{i,q,t}$$

The table presents the results of cybersecurity risk mitigation and informed trading proxies for days preceding earnings announcements (day 0) based on the regression (8). Panel A presents the results of cybersecurity risk mitigation and our primary informed trading proxies leading up to earnings announcements using a 20-day average window. Column (1) presents the results using the natural log of abnormal stock volume. Column (2) presents the results using the natural log of abnormal options volume. Column (3) presents the results using abnormal price range. Column (4) presents the results using abnormal intraday volatility. Panel B reports results of cybersecurity risk mitigation and our primary informed trading proxies leading up to earnings announcements using various one-week windows. Panel B also includes results using additional informed trading proxies . These additional informed trading proxies include the abnormal Amihud's illiquidity measure (*AMI*), abnormal effective spread (*ES*), abnormal quoted spread (*QS*), abnormal realized spread (*RS*), abnormal price impact (*PI*), abnormal order imbalance (*OI*) and abnormal lambda (*Lambda*). Each entry in Panel B reports only the coefficient of β_1 for estimation of the model using the dependent variable for that row and the week represented in the entry's column. All variables are defined in Appendix B. Standard errors are clustered by firm and quarter-year. The sample period is 2012 to 2018. T-statistics are presented in parentheses below the coefficients.