Do firms underreport information on cyber-attacks? Evidence from capital markets

Eli Amir
Tel Aviv University and City University of London
eliamir@post.tau.ac.il

Shai Levi
Tel Aviv University
shailevi@tau.ac.il

Tsafrir Livne
Kenan-Flagler Business School, University of North Carolina
tsafrir@unc.edu

May 2017

Abstract
Firms should disclose information on material cyber-attacks. However, because managers have incentives to withhold negative information, and investors cannot independently discover most cyber-attacks, firms may underreport cyber-attacks. Using data on cyber-attacks that were voluntary disclosed by firms and those that were withheld and later discovered by sources outside the firm, we estimate the extent to which firms withhold information on cyber-attacks. We find withheld cyber-attacks are associated with a decline of approximately 2.6% in equity values in the month they are discovered, and disclosed attacks with a substantially lower decline of 0.6%. The evidence suggests managers do not disclose negative information below a certain threshold, and withhold information on the more severe attacks. Using the market reactions to withheld and disclosed attacks, we estimate that managers disclose information on cyber-attacks when investors already suspect a high likelihood (46%) that an attack has occurred. Overall, our analyses suggest firms underreport cyber-attacks.

Acknowledgement: We thank Eti Einhorn, Tsahi Versano, and seminar participants at the 2017 European Accounting Association annual meeting at Valencia, Hebrew University of Jerusalem, Tel Aviv University, University of Padua for useful comments. We also thank the Blavatnik Interdisciplinary Cyber Research Center, and Henry Crown Institute of Business Research for financial support.

1. Introduction

Firms must disclose cyber-attacks that materially damage their businesses (SEC 2011). However, because investors cannot independently discover most cyber-attacks, and because managers may have incentives to withhold negative unobservable information from investors, firms may underreport cyber-attacks. In this study, we estimate the extent to which publicly traded firms withhold information on cyber-attacks. Specifically, we identify cyber-attacks that were disclosed and attacks that were withheld and later independently discovered. We then use the differential market reaction to these attacks to estimate the extent of underreporting.

Reviewing data on cyber-attacks between 2010 and 2015 suggests many disclosures on cyber-attacks are made after investors discover the attack. Data breaches are revealed to the market, for example, by customers whose information was stolen, or by the hackers themselves.[1] Also, the number of cyber-attacks public companies disclosed, about 300 during that period, seems low in comparison to the thousands of attacks reported by independent sources.[2]

The extent of information withholding is unobservable, and we are aware only of data breaches that are eventually revealed either by the attacked firms or by sources outside the firm. We estimate the extent of withholding from the market reaction to attacks that are revealed, where market reaction serves as a proxy for the damage cyber-attacks cause the firm (e.g., Campbell et al. 2003; Hovav and D'Arcy 2003; Cavusoglu et al. 2004; Kannan et

[1] For example, Target, the US retailer, experienced a data breach involving millions of its customers' credit and debit cards, and after customers and credit card companies revealed the breach, the firm confirmed it. In some cases, the hackers themselves may reveal the breach to the public. For example, hackers breached the LinkedIn network and stole a database containing 6.5 million users' encrypted passwords in June 2013. The hackers later published the attack, hoping to receive help from fellow hackers in cracking these encrypted passwords. After the hackers published the passwords, LinkedIn acknowledged this data breach.

[2] According to Verizon (2015), more than 20,000 data breaches occurred in the US private sector during that time period.

al. 2007; Gordon et al. 2011). We find that in cases in which the firms immediately disclosed the cyber-attack, equity values declined by 0.59% on average in the month of the disclosure. In comparison, the decline in market value was more than four times larger in cases in which the firm withheld the information, and stock prices decreased by 2.59% on average in the month of the discovery of the attack. These findings, which are consistent with Dye (1985), suggest the more severe cyber-attacks are withheld from investors. From the differential market reaction to disclosed and withheld attacks, we estimate that managers disclose cyber-attacks when investors already believe that with a 46% chance, an attack has occurred; when uncertainty about the existence of a cyber-attack is higher, managers withhold the information.

Using alternative estimates of damage caused by cyber-attacks, we also find the more severe attacks are withheld. Specifically, we use damage estimates released by the targeted firms, and also an objective index that measures the severity of cyber-attacks based on type of data breached, the number of records stolen, and the source of the breach. Both damage estimates show, as expected, that more severe attacks are withheld, whereas milder attacks are more likely to be disclosed.

In support of the relation between chances of discovery and withholding, we find withholding firms have lower analyst coverage, weaker corporate governance, and lower litigation risk than disclosing firms. Firms with greater analyst coverage are followed more closely by investors, and the chance of discovery in these firms is higher. Also, firms with stronger governance are less likely to conceal negative news from their investors. Specifically, firms with less entrenched management (Bebchuk et al. 2009), and fewer material weaknesses reported following Section 404 of the Sarbanes-Oxley Act, are more likely to disclose information on cyber-attacks. Using membership in the hi-tech industry as a proxy for high litigation risk, we find disclosing firms are more likely than withholding

firms to be in hi-tech industries. High litigation risk increases the expected loss from withholding information, increasing the attractiveness of disclosure (e.g., Skinner 1994, 1997).

Our study makes a few contributions to the literature. First, we use disclosure theory to explain the market effects of cyber-attacks. Prior studies that examine the stock price reaction to cyber-attacks find mixed results. For instance, Cavusoglu et al. (2004) find data breaches have a statistically significant negative effect on stock prices. By contrast, Campbell et al. (2003) and Kannan et al. (2007) find the market effect of breaches is generally insignificant.[3] Gordon et al. (2011) report a decrease in the effect of breaches on stock prices over time. They conjecture that with increased media reporting of data breaches without apparent devastating effects on targeted corporations, investors lowered their assessment of the costs of data breaches. Kvochko and Pant (2015) also review recent cases in which large data breaches had a small impact on stock prices. Hilary, Segal, and Zhang (2016) find that the average market reaction to data breaches between 2005 and 2014 was not different from zero statistically. Consistent with the latter studies, we find the negative reaction to most breaches in our sample (2010-2015) is quite small. However, unlike prior studies, we distinguish between cyber-attacks that were voluntary disclosed and those that were withheld from investors and later independently discovered, and find that in the latter cases the market reaction is negative and significant. These results suggest cyber-attacks that are unknown to investors are more likely to be severe, and that the market reaction reported in prior studies understates the damage cyber-attacks cause firms.

We also contribute to the empirical disclosure literature. Prior literature examines the different timing of good- and bad-news disclosures. For example, Kasznik and Lev (1995)

[3] Mixed results exist also for specific types of data breaches. For example, Hovav and D'arcy (2003) and Kannan et al. (2007) find denial-of-service attacks have an insignificant effect, whereas Ettredge and Richardon (2003) find this kind of attack has a significant negative impact on the market value of firms.

examine if firms warn investors of upcoming negative earnings surprises. Amir and Ziv (1997) find firms delay the adoption of new accounting standards with negative financial effects. Chambers and Penman (1984) find late earnings announcements contain, on average, worse news than early announcements. Kothari et al. (2009) find the magnitude of negative stock price reaction to bad news is greater than the magnitude of positive stock price reaction to good news, and infer from their evidence that managers accumulate and withhold bad news up to a certain threshold, but leak and immediately reveal good news. In the case of cyber-attacks, however, the withheld information will likely never be revealed to investors. Also, for the breaches that are eventually revealed, the data indicate when the firm first learned of the attack, and therefore whether information withholding occurred. This setting and data enable us to clearly distinguish between cases of disclosing and withholding, and show that, consistent with theory, managers withhold more negative information and voluntarily disclose less severe attacks. This setting also allows us to examine the different market reactions to withholding and disclosure, and to estimate when withholding information is worthwhile for managers. Using market reactions to withheld and disclosed attacks, we show that managers disclose cyber-attacks only when the likelihood that investors believe an attack is imminent is high.

Our findings are consistent with underreporting of cyber-attacks. If regulators wish to ensure information on cyber-attacks reaches investors, they should consider tightening disclosure requirements of cyber-attacks. The US Congress has recently asked the SEC to revise cyber-security disclosures, and the SEC is currently considering measures that would require publicly owned companies to disclose more information about their cyber-security and data breaches.[4]

[4] See http://thehill.com/policy/cybersecurity/229431-sec-weighs-cybersecurity-disclosure-rules.

2. Hypothesis development

Cyber-attacks are often unobservable to the public when they occur, and even when investors know about the attack, prior literature finds the average damage is small (e.g., Campbell et al. 2003; Kannan et al. 2007). Investors' uncertainty about the existence of attacks and the low average damage provide targeted firms with the opportunity to withhold negative information about cyber-attacks. When managers are known to withhold information, investors will discount the stock to reflect the worst possible news, which will drive managers to make full disclosure (Grossman and Hart 1980; Grossman 1981). However, when investors are uncertain about whether managers possess negative information, withholding is possible (Dye 1985). In the absence of disclosure, investors will reduce stock prices only by the expected value of the bad news, which in the case of cyber-attacks equals the probability that an attack occurred (and the manager is withholding the information), times the average damage. Because the damage of cyber-attacks is on average small, the expected loss from not disclosing is low, and withholding is an attractive option for managers.

We use a setting similar to that used by Dye (1985) to develop our main hypothesis—firms will withhold information on the more severe cyber-attacks and voluntarily disclose the milder ones.[5] Assume a cyber-attack on the firm with a probability p and a loss x. Only the manager learns of the attack and the damage, x. Managers will withhold information on the damage, x, when the loss from disclosing (which is equal to x) is higher than the expected loss from withholding. Because investors know the ex-ante distribution of the damage, $\tilde{x}$, and the probability of cyber-attacks in the case of no disclosure, they can

[5] Dye (1985) assumes firm owners wish to maximize current share price and provide managers with incentives to withhold negative information. The assumption that, in general, managers wish to maximize share prices is reasonable because their career and reputation are usually linked to share prices.

estimate the likelihood that the decision not to disclose is due to withholding, *prob(withholding)*, and the expected loss in the case of withholding, which is $prob(withholding)E(\tilde{x}|withholding)$.[6] Therefore, managers will withhold information if $prob(withholding)E(\tilde{x}|withholding) > x$. Using this setting, Dye (1985) shows a disclosure threshold, $\underline{x}$, exists above which managers will disclose information, which equals $prob(withholding)E(\tilde{x}|\tilde{x} < \underline{x}) = \underline{x}$. The disclosure threshold, $\underline{x}$, equals the probability, *prob(withholding)*, that the manager has bad news and withholds it, times the expected value of the bad news withheld. Managers withhold bad news if the damage from the attack is lower than $\underline{x}$. It follows that managers will withhold information on the more severe cyber-attacks—the ones that will cause a loss in stock prices that is below the disclosure threshold, $\underline{x}$.

Clearly, the probability of discovery by investors affects disclosure. Because the expected value of the bad news withheld, $E(\tilde{x}|\tilde{x} < \underline{x})$, is negative, when investors believe the probability of managers holding negative information, *prob(withholding)*, is high, the disclosure threshold, $\underline{x}$, will be lower and managers will disclose more negative news.[7] Using market reactions, we empirically estimate the probability of withholding, *prob(withholding)*, at which managers chose to disclose a cyber-attack.

The probability of withholding is equal to $\frac{\underline{x}}{E(\tilde{x}|\tilde{x} < \underline{x})}$. To estimate this probability, we need empirical proxies for the disclosure threshold, $\underline{x}$, and the expected value of bad news withheld, $E(\tilde{x}|\tilde{x} < \underline{x})$. We use the average stock returns in the withholding cases that are

[6] Dye (1985) and Jung and Kwon (1988) differently calculate the likelihood of withholding conditioned on no disclosure, *prob(withholding)*, in this setting. Either way, the likelihood of withholding will be a function of the probability of a cyber-attack and the probability that the damage will be below the disclosure threshold and therefore withheld by managers.

[7] Jung and Kwon (1988) show how, in this setting, an increase in the probability with which investors believe managers have negative information will lower the disclosure threshold and will trigger the release of information managers would otherwise withhold.

discovered by investors as a proxy for the expected value of bad news withheld, $E(\tilde{x}|\tilde{x} < \underline{x})$.[8] We assume the average damage of the discovered attacks is representative of the damage in withheld cases. To estimate the disclosure threshold, $\underline{x}$, we use the average return reaction in the cases in which managers disclosed the cyber-attack. However, managers disclose losses whenever they are low enough (above the threshold), and we observe the average loss. To estimate the threshold loss, we assume, similar to Dye (1985), that the loss is uniformly distributed on the interval [$\underline{x}$,0], where the threshold $\underline{x}$ is a negative number.[9] The expected loss disclosed is hence $\frac{\underline{x}}{2}$. It follows that the probability that managers are withholding bad news on cyber-attacks is

$$prob(withholding) = \frac{2*Return\ reaction\ to\ immidiate\ disclosure}{Return\ reaction\ to\ discovery\ of\ withholding}. \quad (1)$$

As we show below, the average market reaction is, for example, -0.59% in the month following an immediate disclosure of the breach by firms, and -2.59% when the breach was not disclosed, but investors later discovered it. These estimates imply managers disclose cyber-attacks when investors already believe that with a probability of 46%, an attack has occurred.

3. Data

Our first data source is the AuditAnalytics cyber-attacks database, which documents 186 incidents between 2010 and 2015. For 162 of these incidents, we obtain stock returns from the Center for Research in Security Prices (CRSP). The second data source is the

[8] As discussed below, the average market reaction to attacks that are discovered may be a biased estimate of the damage. Specifically, the decrease in price upon discovery may be larger than the damage due to the negative reputation effects and litigation risk associated with withholding. In this case, our withholding-probability estimate will be downward biased.

[9] Dye (1985) uses the same assumption in the illustrative example of his theorem (p. 129). As discussed below, even if the loss is not uniformly distributed, we can still estimate the minimal probability of withholding, because the disclosure threshold will not be higher than the actual return reaction in the cases in which firms disclosed the cyber-attack.

VCDB VERIS community database, which contains thousands of documented incidents, of which only a small fraction relates to public companies. Description of the VCDB VERIS database can be found, for example, in the Verizon (2015) Data Breach Investigations Report. According to the report, the database includes information on data breaches collected by Verizon during its "paid external forensic investigation" services, and by 70 other cyber-security companies and organizations. The identity of most firms on the database is unknown, and we find an additional 158 data breaches of public firms between 2010 and 2015 that are not included in the AuditAnalytics' database.

The anonymity of breaches balances the needs of VCDB VERIS community in data, and firms' privacy. Most private firms do not share information on their operations with third parties, and many of the anonymous records are likely of private firms. Public firms will also be reluctant to reveal negative information to competitors and investors, so a large number of anonymous breaches in the database can therefore contain records of public firms. From both data sources together, we obtain data for 320 incidents involving 180 publicly traded companies between 2010 and 2015, of which 66 firms had more than one cyber-attack.

For descriptive purposes, we classify cyber-attacks into three categories following Gordon et al. (2011).[10] The first category -- "availability" -- includes breaches that stop the business from making its services available to customers (also known as denial of service). We include in this category cases in which the breach can jeopardize availability. For example, cases in which hackers gain access and can disrupt main systems, or steal intellectual property. The second category -- "confidentiality" – are breaches that allow unauthorized users access to confidential information as bank account credentials, credit card data, medical records, or insurance history. The third category includes breaches of

[10] In section 4.4 below, we control for the effect of attack of type on announcement returns.

information “integrity”, i.e. breaches that compromise the reliability of a database or a website. We include in this category cases linked to stealing email addresses, usernames, or passwords to social media. We classify all breaches in the sample into one of the three categories, except for three breaches for which the information is not available in the data and the attack type is unknown. Table 1 summarizes information on our sample by year and attack type.

(Table 1 about here)

4. Results

4.1 Firm Characteristics

We classify sample firms into three groups according to their disclosure policy. We classify a cyber-attack incident as “Disclosing” if the firm disclosed the cyber-attack before an outside party discovered it (72 disclosure cases, 22.5%). We also classify cyber-attacks as "Disclosing" when the attack is concurrently discovered by outsiders and disclosed by the firm (91 cases, 28.4%). We classify a cyber-attack as "Withholding" if the firm had not disclosed the cyber-attack for at least two days after it learned of its occurrence, and a party outside the firm consequently discovered the attack (57 cases, 17.8%).[11] We believe that disclosing the information within a day or two of the breach does not amount to withholding, and does not expose the firm to litigation and decrease in reputation, which are associated with withholding of negative information.[12]

[11] We classify cases as “Withholding” only cases in which it is clear that the firm learned of the attack before a party outside the firm discovered the attack. In many cases, firms eventually disclose the date on which they learned of the attack, and this date is provided by the data vendor, AuditAnalytics, and is collected for VCDB VERIS data cases.

[12] In a sensitivity analysis, when we define Withholding if the firm had not disclosed the cyber-attack for at least a week, we find stronger results. For example, the returns in the month of discovery, Ret(-1, 20) that are reported in Table 4 are -3.57% for these Withholding cases.

Many of the attacks that are discovered by outside parties turn out to be immaterial according to the attacked firm. In 62 cases, we find that after discovery of the breach, the firm clarified it was immaterial. In 38 cases, the firm did not respond to the cyber-attacks outsiders discovered, and hence implicitly communicated the event was immaterial.[13] Hence, we classify these 100 cases (31.3%) as "Immaterial." In Table 2, we present descriptive statistics on the three subsamples and examine whether the subsamples are materially different from each other. Panel A of Table 2 compares Disclosing to Withholding; Panel B compares Disclosing to Immaterial.

Firms are more likely to disclose cyber-attacks when the likelihood of discovery of the breach by outside parties is higher. As a measure of outside monitoring by investors, we use the number of analysts following the firm. Firms with greater analyst coverage are followed more closely by investors, and these firms are more likely to disclose negative information, such as data breaches. We measure analyst coverage ("Analysts") as the number of analysts on I/B/E/S during the year. As Panel A shows, firms that disclosed cyber-attacks are followed by 13.30 analysts, on average, whereas withholding firms are followed by 9.11 analysts, on average, and the difference is statistically significant at the 0.01 level. Panel B of Table 2 shows the average number of analysts following firms with Immaterial cyber-attacks was 16.25, which is larger than the number of analysts following Disclosing firms, at the 0.05 level. As we expected, Withholding firms are, on average, followed by fewer analysts, which is consistent with less monitoring by investors and hence the lower probability of independent discovery.

Firms with stronger corporate governance are less likely to withhold negative news from their investors, because stronger governance is associated with stronger fiduciary

[13] In case of litigation, for example, a non-response will be considered a (passive) statement that the event was immaterial.

responsibility to the firm's owners. As a measure of governance strength, we use the number of material weaknesses the firms reported under Section 404 of the Sarbanes-Oxley Act of 2002 in the five years preceding the breach (SOX404). Section 404 requires all publicly traded companies to establish internal controls and procedures for financial reporting, and its purpose is to reduce the possibility of corporate fraud. Material weaknesses are reported when there are deficiencies in controls that create a reasonable possibility of misstatements in the firm's financial statements (e.g., Ge and McVay 2005). Although a material-weakness report does not mean a material misstatement has occurred, it means internal controls may not be strong enough to detect or prevent a material misstatement on a timely basis. However, the existence of material weaknesses in controls increases the likelihood that firms withhold information on losses associated with cyber-attacks. Data on material weaknesses are available on the AuditAnalytics database.

As Panel A of Table 2 shows, Withholding firms had more material weaknesses than Disclosing firms in the years preceding the withholding. The average of SOX404 is 0.56 and 0.07 for Withholding and Disclosing firms, respectively, and the difference is statistically significant at the 0.01 level. As Panel B shows, no difference exists between the average SOX404 of Disclosing firms and that of firms that experienced immaterial cyber-attacks.

We also use Bebchuk et al.'s (2009) entrenchment index as a governance metric. Larger index values suggest weaker corporate governance. As Panel A of Table 2 shows, Disclosing firms have lower entrenchment-index values (average 1.40) than Withholding firms (average 1.74), and the difference is significant at the 0.04 level. According to Panel B, the entrenchment index of Disclosing firms is higher than that of firms with Immaterial damage (average 1.40 vs. 1.27, respectively, significant at the 0.08 level). According to the Bebchuk et al.'s (2009) entrenchment index, firms that disclose information on cyber-attacks have stronger corporate governance than firms that withhold information on cyber-attacks.

We expect that firms with higher litigation risk will disclose information on cyber-attacks (Skinner, 1994). Similar to Kasznik and Lev (1995), we use membership in hi-tech industries as a proxy for high litigation risk. We use a Hi-Tech indicator that equals 1 for firms in Drugs (SIC codes 2833-2836), R&D Services (8731-8734), Programming (7371-7379), Computers (3570-3577), and Electronics (3600-3674); and 0 otherwise. As Panel A of Table 2 shows, 21% of Disclosing firms are in these hi-tech industries, whereas only 7% of Withholding firms are in the hi-tech sectors (difference is statistically significant at the 0.01 level). As Panel B shows, the percentage of firms with immaterial cyber-attacks that are in the hi-tech sectors (22%) is similar to the percentage of hi-tech firms in the Disclosing subsample.

Withholding is more likely when the damage of the cyber-attack is larger. We use two measures of the cost of the cyber-attack. "Damage" is an estimate of the damage the cyber-attack caused, as provided by the targeted firm, divided by market value of equity at the beginning of year. We obtained a total of 40 such damage estimates for our sample, which were provided by the targeted firms in their financial statements subsequent to the attack. The second variable, "Severity," is an index taking values from 0 to 10, depending on the severity of the cyber-attack (0=low damage and 10=very high damage). Gemalto (an international digital security company) created the index to measure the severity of cyber-attacks, and the index can be calculated for the entire sample. The index rates the severity of data breaches based on the type of data breached, the number of records stolen, the source of the breach, and whether or not the hackers used the data stolen. Severity and Damage are highly correlated (Pearson correlation of 0.49).

Consistent with our hypothesis, Table 2 shows the damage in the Withholding cases (0.018 or 1.8% of the market value of equity) is larger, on average, than that in the Disclosing cases (0.6% of the market value of equity). The difference is significant at the

0.10 level. Also, as expected, both Damage and Severity are higher for Disclosing cases than for Immaterial cases.

Table 2 also presents the profitability (ROA) in the year before the attack, measured as net income divided by total assets. Firms may time the disclosure of negative information based on their overall profitability. For instance, firms withhold the negative news in good years, and clean the slate and disclose the negative information in periods with weaker profitability (Levitt, 1998). As the table shows, Disclosing and Withholding firms report similar ROAs, and the differences between the groups are not statistically significant. Therefore, differences in profitability do not explain the decision of firms to disclose or withhold information on cyber-attacks.

Finally, we examine whether the three subsamples differ from each other in terms of firm size, measured as the market value of equity at the beginning of the year (MV). In line with the findings on the number of analysts, we find that Disclosing firms have larger market values than Withholding firms; larger firms are often followed by more analysts. Also, firms with Immaterial cyber-attacks have larger market values than Disclosing firms.

(Table 2 about here)

Table 3 presents a multivariate logistic regression for the variables presented in Table 2. The purpose of this analysis is to examine which of the variables have a significant marginal effect. The dependent variable obtains the value 1 for Disclosing firms, and 0 otherwise. As Model 1 shows, only four of the variables are significant at the 0.10 level or better. Consistent with our hypothesis, the severity of the attacks that are withheld is, on average, larger than the severity of those attacks that are immediately disclosed (p-value = 0.06). Second, poor corporate governance metrics, that is, higher SOX404 and Entrenchment, are associated with less disclosing and more withholding of information on cyber-attacks (p-value of 0.03 and 0.10, respectively). This result is consistent with the claim

that stronger corporate governance leads to more timely disclosure of negative information. Finally, membership in hi-tech industries, which serves as a proxy for higher litigation risk, is positively associated with disclosure (p-value = 0.03). In model 2, we compare Disclosing cases with cases with immaterial cyber-attacks. As expected, we find the severity index is higher for Disclosing than for Immaterial cases (at the 0.01 level). We also find Immaterial cases are reported by larger firms (at the 0.01 level) than Disclosing cases. Overall, the analysis in Table 3 suggests stronger governance and higher litigation risk are associated with more disclosure and less withholding of information on cyber-attacks, and the severity of withheld cyber-attacks is larger than the severity of disclosed cyber-attacks.

(Table 3 about here)

4.2 Market Reaction to Cyber Attacks

Table 4 shows the cumulative risk-adjusted returns surrounding the date a cyber-attack became known to investors. In the main tests, we adjust the returns for risk using the value-weighted market return reported by CRSP (variable VWRETD). Specifically, we compute the difference between the buy-and-hold returns of the stock and the buy-and-hold-returns of the value-weighted market portfolio.[14] This simple risk adjustment can be applied to all 320 data breaches in our sample.[15]

We present return reaction for three return windows: (i) from one trading day prior to the disclosure until three trading days after the disclosure, denoted as Ret(-1,3); (ii) from one trading day prior to the disclosure until 10 trading days after the disclosure, denoted Ret(-1,10); and (iii) from one trading day prior to the disclosure until 20 trading days after the disclosure, denoted Ret(-1,20), which is approximately a calendar month.

[14] This approach is equivalent to using a beta equal to 1. Because firm-specific beta estimates are noisy (Fama and French, 1996), we use this approach for simplicity and to maximize sample size.
[15] Using alternative risk adjustments for smaller samples, we find similar results. See Table 7.

Focusing on the first and shortest return window, the average market reaction to Disclosing is -0.17% but is not statistically different from zero. This result suggests data breaches that firms disclosed did not have a significant marginal effect on the stock value. These findings coincide with negative and insignificant stock returns that other studies find around data breaches (e.g., Campbell et al. 2003; Kannan et al. 2007). By contrast, we find the average market reaction to the 57 cases in which firms withheld information on cyber-attacks is -1.19% (significant at the 0.01 level); that is, stock value decreased 1.19% from one day prior to disclosure until three days after investors independently discovered the breach. In addition, in cases of Immaterial cyber-attacks (100 cases), we find the average market reaction is 0.04% but not statistically significant.

Within 10 days of the discovery date, stock prices continued to decline for Withholding firms. Specifically, returns 10 and 20 trading days after discovery were, on average, -1.92%, and -2.59% (both are significant at the 0.01 level), respectively. The cumulative risk-adjusted returns between trading days 4 and 20 are -1.42% (significant at the 0.02 level) for Withholding firms. These results suggest investors take a few days to understand the firm withheld material negative information, and to fully respond to the information.

In sum, results support the hypothesis that news on withheld cyber-attacks is more negative than news on disclosed cyber-attacks. Consistent with prior studies, cyber-attacks in general have a low negative effect on the market value of equity; however, we find that in cases in which firms withheld information from the public and investors eventually revealed the breach, the market reaction was negative and significant. The findings are consistent with our hypothesis that firms withhold negative information below a certain threshold, disclose information on less severe cyber-attacks, and keep from investors more severe cyber-attacks that can significantly affect stock prices.

(Table 4 about here)

Figure 1 presents the cumulative risk-adjusted returns from one trading day prior to the discovery date until 20 days after the discovery date for three groups: Withholding, Disclosing, and Immaterial cyber-attack cases. The results show a clear pattern: the stock price decrease in the Withholding cases is larger than in cases in which firms immediately disclose the breach to investors. The negative reaction to withholding information is not temporary; it persists long after the cyber-attack is discovered. Going beyond the 20 trading days presented in the graph, we find that returns of the Withholding portfolio are -3.56% (p-value < 0.01) after 30 trading day, and -4.14% (p-value of = 0.01) after 60 trading days. In comparison, cumulative returns of the Disclosing portfolio become positive after 30 trading days, 0.17% (p-value = 0.41), and remain positive after 60 trading days, 0.73% (p-value = 0.26). Returns in longer windows will be driven not only by the data breach, but also by other events, and the power of the test will therefore be lower, especially in small samples like ours.

(Figure 1 about here)

If earnings are announced during the 20 days after the cyber-attack is discovered, the earnings announcements and not the cyber-attack could affect the results. We therefore exclude from the return calculation the three days around the announcement, from a day before to a day after the quarterly earnings announcement.[16] We find similar results to those presented in Table 4. For example, Ret(-1,20) in the Withholding cases is -1.93% (significant at the 0.01 level); in Disclosing cases, it is -0.43% (p-value of 0.23); and in the Immaterial cases, it is 0.45 (p-value of 0.19).

[16] Less than 8% of the attack-discovery dates exactly coincide with the earnings-announcement date, and when excluding these observations, we get similar results.

We also plot in Figure 2 the risk-adjusted returns for Disclosing, Withholding, and Immaterial groups from one trading day before to 20 trading days after quarterly earnings announcements in the year of the cyber-attack. If earnings announcements drive the reaction in Figure 1, the plot in Figure 2 should exhibit similar patterns to those in Figure 1. The results of this analysis show the returns 20 trading days after the earnings announcement is, on average, -0.40% (p-value of 0.50) for the Disclosing firms, and is 0.62% (p-value of 0.50) for the Withholding firms, and the difference between these average returns is not statistically significant. These results suggest that reaction to earnings announcements of Disclosing and Withholding firms was not different during the year of the cyber-attacks. The different return reactions to cyber-attack news of Withholding and Disclosing firms cannot be attributed to differential earnings performance, but to the cyber-attack and its corresponding damage and disclosure policy.

4.3 Implied probability of withholding

We estimate the implied probability of cyber-attack withholding using eq. (1) and the market reaction to disclosing and withholding information on cyber-attacks. Panel B of Table 4 shows the results. Based on the return reaction in the three days after the discovery date, Ret(-1,3), we estimate the probability of withholding to be about 30%, which is twice the return reaction of -0.17% in Disclosing cases, divided by the return reaction of -1.19% in withholding cases. If these return reactions indeed capture the damage caused by the cyber-attack, managers will disclose the cyber-attack only when investors believe the probability that managers hold negative information is higher than 30%.

When we use a wider return window after the discovery date, we get higher estimates of the probability of withholding. After the discovery date, more information about the breach and its damage flow to the market, and the market revises its cost estimates. As Panel B shows, cost estimates that are based on the 10-day return window imply the probability of

withholding is 47%, and those based on the 20-day return window suggests the probability of withholding is 46%. That is, only when the chance that investors already know of the cyber-attack is at least 46% do firms choose to disclose the information.

As discussed in section 2 above, our measure of the implied probability of withholding assumes the damage is uniformly distributed, and therefore the disclosure threshold in disclosure cases is estimated to be twice the average returns. The distribution of loss may be different, but in any case the disclosure threshold will not be higher than the actual returns in the Disclosing cases. Therefore, at the minimum the implied probability of withholding is 23% according to the 20-days window. Our estimate of the loss in withholding cases may be also biased. We assume the average returns in the withholding cases that are discovered are representative of the damage. However, empirically, the decrease in price upon discovery may be larger than the damage because of negative reputation effects and litigation risk that can be associated with withholding, in which case, our withholding-probability estimates are downward biased. Another assumption we make is that the return reaction to the cyber-attack starts on the discovery date. To validate this assumption, we check and find that the cumulative risk-adjusted returns between day -10 and day -2 before the discovery date is 0.91%. If investors had started suspecting managers were withholding negative information, prices would have declined before the discovery date.

To estimate the statistical significance of the withholding-probability estimates, and specifically that the withholding probability triggering disclosure is higher than zero, we assume it is a proportion that is distributed between 0 and 1 for a sample 163 observations. We find the withholding probability estimates are greater than zero at least at the 0.01 level. Results are similar when we use bootstrapping (e.g., Chernick 2007) and use 100 random samples with replacement from the original sample of return reactions to estimate the standard deviation.

Next, we examine whether the results in Table 4 reflect the market reaction to the disclosure decision after controlling for the damage caused by the cyber-attack. We hypothesize that firms sustaining more damage are more likely to withhold the breach information from investors, whereas firms sustaining less damage voluntarily disclose the breach.

We use damage estimates disclosed by the attacked firms. In our sample, we find that 40 firms reported the dollar value of the damage caused by the cyber-attack in a press release or in subsequent financial statements. To control for the effect of the damage on the return reaction to the cyber-attacks, we use the following OLS regression:

$$\mathrm{Ret}(-1,3)_i = \beta_0 + \beta_1 Disclosing_i + \beta_2 Withholding_i + \beta_3 Damage_i + \varepsilon_i, \quad (2)$$

where Ret(-1,3) is the cumulative risk-adjusted returns from one day prior to discovery until three days after discovery, Disclosing is an indicator variable that equals 1 for disclosing cases, Withholding is an indicator variable that equals 1 for withholding cases, and Damage is the damage estimate provided by the firm, divided by the market value at the beginning of year. We expect the slope coefficient on Withholding to be negative. Table 5 presents the estimation results.

We find that after controlling for damage, the return reaction to withholding is negative and significant upon the discovery of the attack. As model 2 shows, the coefficient on Withholding is -0.023 (significant at the 0.10 level); the coefficient on Damage is -0.382 (significant at the 0.06 level). Also, the coefficient on Disclosing is -0.009 and statistically insignificant. The results suggest the equity values of withholding firms decreases beyond the damage estimates provided by their managers after the attack. Uncertainty exists regarding the damage caused by cyber-attacks, and managers may use this uncertainty to provide low damage estimates. Additionally, a decrease in value may be driven by other information withholding conveys. If investors eventually learn of the cyber-attack from other

sources, they are likely to update their beliefs on the integrity and quality of management. Whether managers withheld information, or just failed to monitor their information systems and identify the attack, investors will take the lack of timely disclosure as a negative sign. Additionally, firms that withhold bad news may face litigation once it is discovered (e.g., Skinner 1994, 1997; Lev and Kasznik 1995), which will also negatively affect equity value.

Because the regression results suggest that decrease in price upon discovery of withholding is partly driven by the negative reputation effects associated with withholding, the withholding-probability estimates based on returns will be downward biased. For example, when we use the dollar-damage figures that are reported by firms (see Table 2) instead of the market reaction to calculate the probability of withholding, we get that, following eq. (1), the probability of withholding is 66% (2*0.006/0.018).

(Table 5 about here)

Next, we examine the sensitivity of our results to the inclusion of the firm-characteristic variables introduced in Table 2. The purpose of this analysis is to alleviate concerns that certain firm characteristics make certain firms more vulnerable to cyber-attacks and hence more likely to be included in the sample. We use the following regression model:

$$\begin{aligned} \mathrm{Ret}(-1,3)_i &= \delta_0 + \delta_1 Disclosing_i + \delta_2 Withholding_i + \delta_3 Severity_i \\ &+ \delta_4 Analysts_i + \delta_5 HiTech_i + \delta_6 SOX404_i + \delta_7 Entrenchment_i \\ &+ \delta_8 ROA_i + \delta_9 LogMV_i + \varepsilon_i \end{aligned} \quad (3)$$

.

The results, which are reported in Table 6, show the coefficient on Withholding is negative and significant in most specifications. The coefficient on Disclosing is negative and statistically insignificant, indicating the returns in the disclosing cases are not different than in the immaterial-damage cases. The coefficient on Severity is also negative, as expected, and statistically significant at the 0.10 level or better. Severity is a proxy for the damage, and

the low significance of the coefficient relative to the coefficient on the actual-damage variable used in Table 5 may be attributed to the noisy nature of this proxy.

Overall, the results in Tables 4 and 6 are similar. For example, the average return reaction in the three days after discovery of firms that withheld information on the cyber-attack is 1.3%-1.4% lower than the return reaction to cyber-attacks on other firms.

(Table 6 about here)

4.4 Sensitivity Analyses

First, we control for the effect of endogeneity. The severity of the attack can affect firms' decision to disclose or withhold, and this endogeneity may bias the regression results presented in the above analysis. Before detailing the empirical steps we take to deal with this concern, let us first emphasize that endogeneity is part of the hypothesis. We hypothesize that firms will withhold larger cyber-attacks and disclose smaller one. Put simply, we try to prove that endogeneity between the severity of attack and disclosure decision exists. In the univariate analysis, for example in Tables 1 and 3, this endogeneity is (econometrically) not a problem. In the regression analysis, the endogeneity may create a correlated omitted variable problem, and bias the coefficients. To deal with this problem, we use an instrumental variable estimation.

We use the state of incorporation as an instrumental variable. The state of incorporation affects the decision to disclose the cyber-attack, but is not associated with the severity of the attack, and therefore can serve as an instrument. Some US states require firms to disclose attacks regardless of their severity. For example, California firms need to notify customers or individuals whose private information was breached, and if the number of individuals that were affected is greater than 500, then the company needs to also notify the Attorney General of California (Cal. Civ. Code § 1798.82). Therefore, incorporation in California is expected to affect the disclosure decision to disclose, but not the damage—i.e.,

the incorporation in California will not ex-ante bring about more severe attacks on firms. This fact allows us to use the state of incorporation as an instrumental variable.

We define 26 states in which firms are required to notify the state attorney general of certain breaches has high-disclosure states. The high-disclosure states include CA, CT, FA, HI, IN, IA, LA, ME, MD, MA, MO, MT, NE, NH, NJ, NY, NC, ND, OK, OR, RI, SC, VT, VA, WA, PR, and HDState is an indicator variable that equals 1 for these states, and 0 for other states of incorporation. We use the following 2SLS estimation:

$$Withholding_{it} = \alpha + \beta HDState_{it} + \varepsilon_{it} \quad (4a)$$

$$Ret(-1,3)_{it} = \alpha + \beta_1 \overline{Withholding_{it}} + \beta_2 Severity_{it} + Controls + \theta_{it}, \quad (4b)$$

where the variables are similar to those in eq. (3) above. In the first stage, we estimate (4a), and use the expected value, $\overline{Withholding}_{it}$, in the second stage for estimation of (4b).

First note that of the 51 cyber-attacks against firms incorporated in high-disclosure states, only 11.8% were withheld by the firms, versus 19.0% of the 269 attacks against firms incorporated in other states. Estimating Eq. (4a) with year fixed effects, we find that the coefficient on HDState is -0.075, and lower than zero at the 0.01 level. This result suggest that withholding is lower in high-disclosure states. Using the expected value from estimation of eq. (4a), we estimate (4b) and present the estimation results in Table 7.

Table 7 shows the coefficient on withholding is negative, -0.012, based on OLS regression. However, once 2SLS is used to control for endogeneity, the coefficient on $\overline{Withholding}_{it}$ is positive, 0.099. The coefficient on Severity is still negative and significant. Together the results suggest that withholding itself, e.g. the negative reputation that may be associated with withholding, does not negatively effect on announcement returns, and only the greater severity of the cyber-attacks withheld lead to the negative announcement returns.

(Table 7 about here)

Second, we correct for a possible bias due to self-selection, the fact that not all withheld attacks are being ex-post discovered. In the main analysis we assume that discovered attacks represent the population of withheld attacks. However, if discovered attacks are a biased sample of withheld attacks, then the estimation we present above can be biased. We employ Heckman's procedure to correct for this possible bias. Specifically, we use the following estimation:

$$\begin{aligned} Withholding_i = {} & \delta_0 + \delta_1 HDState_i + \delta_2 Severity_i + \delta_3 Analysts_i + \delta_4 HiTech_i \\ & + \delta_6 SOX\,404_i + \delta_7 Entrenchment_i + \delta_8 ROA_i + \delta_9 LogMV_i + \varepsilon_i \end{aligned} \quad (5a)$$

$$Ret(-1,3)_{it} = \alpha + \beta_1 Withholding_{it} + \beta_2 Severity_{it} + Controls + MR + \theta_{it} \quad (5b)$$

Eq. (5a) explains the selection of Withholding sample. As discussed above, firms incorporated in high-disclosure states are less likely to withhold cyberattack. Higher severity attacks are more likely to be withheld. Firms with higher analyst coverage, hi-tech firms, and firms with better corporate governance are less likely to withhold information on cyberattacks. We calculate the inversed Mills ratio (MR) based on the estimation of (5a), and use it to correct the possible bias due to selection bias in (5b). As the results in Table 8 show, the estimation results of eq. (5b) are similar to those presented in the main analysis. Specifically, we find that the coefficients on Withholding and Severity are negative and significant.

(Table 8 about here)

In the main tests, we adjust stock returns for risk using the value-weighted market return. This simple risk adjustment can be applied to all 320 data breaches in our sample, and we use it to maximize the sample size. In Table 9, we present the results with returns adjusted for risk, using Daniel et al.'s (1997) size, book-to-market, and momentum quintile portfolios (computed for 240 breaches) and CRPS size-decile portfolios (computed for 277 breaches). These two alternative risk-adjusted returns yield similar results.

(Table 9 about here)

To demonstrate the market reaction to news announcements of withholding and disclosing firms in general does not differ, we estimate eq. (3) for earnings-announcement days. As Table 10 shows, the coefficients on the Withholding and Disclosing dummy variables are not different from zero, suggesting the effect of the cyber-attack and not a general earnings effect drives the different market reaction we record for disclosing and withholding firms.

(Table 10 about here)

Our results are similar when we control for the type of attack. Using the Gordon et al.'s (2011), classification we add three indicator variables to eq. (3)—one for each attack type (Availability, Integrity, and Confidentiality). The main results (not tabulated) are similar to those reported in Table 6. Specifically, the coefficient on Withholding is -0.016 (p-value < 0.01), and the coefficient on Disclosing is -0.005 (p-value = 0.20). Also, none of the attack-type indicators is significantly different from zero (p-values range from 0.43 to 0.47).[17] These results suggest the type of attack does not drive the effect of withholding on the market reaction.

We perform additional robustness tests to rule out alternative explanations to our results. First, to rule out the possibility that market-wide effects are driving our results, we perform the analysis using raw returns instead of market-adjusted returns. For example, we may underestimate the cost of the cyber-attack due to information spillover (e.g., a discovery of an attack on one firm may have an effect on other firms), and by subtracting market returns, we underestimate the cost of the cyber-attack. Using raw returns as the dependent variable in eq. (3) yields similar results (not tabulated for brevity). The coefficient on

[17] On a univariate level, availability, confidentiality, and integrity attacks are associated with returns, Ret(-1,3), of -0.64%, -0.33%, and 0.01%, respectively. Gordon et al. (2011) similarly find that availability attacks are associated with larger damages than confidentiality attacks, and integrity attacks are associated with the lowest damages. Once we control for the damage, the attack type does not provide any additional explanatory power.

Withholding is -0.016 and significant (p-value of 0.02), and the coefficient on Disclosing is -0.007 and statistically insignificant (p-value of 0.18). The results indicate market-wide factors do not drive the different effect of disclosing and withholding on market reaction.

Second, several firms experienced multiple cyber-attacks, and investors may be reacting differently to a first cyber-attack than to a second or third attack. To control for multiple attacks, we add indicator variables for the first attack on firms, an indicator variable for the second attack on firms, and so on. Our results and inferences do not change. The coefficient on Withholding is -0.015 and significant (p-value of 0.01), and the coefficient on Disclosing is -0.005 and statistically insignificant (p-value of 0.20).

Finally, Gordon et al. (2011) report a decline in the market reaction to cyber-attacks over time. To examine whether change in market reaction over our sample period drives the results, we estimate eq. (3) with a time-ordinal variable. The coefficient on this variable is not statistically different from zero (-0.0001, p-value = 0.47). Also, the results do not change: the coefficient on Withholding is -0.016 (p-value < 0.01), and the coefficient on Disclosing is -0.005 (p-value = 0.22). The results of this sensitivity analysis suggest change in market reaction over time does not drive the different effect of disclosing and withholding on market reaction that we report.

6. Conclusion

Cyber-attacks are currently considered one of the main risks firms must manage. Prior studies raised doubts on whether cyber-attacks are indeed so harmful. In particular, studies used the market reaction to cyber-attacks to show that the loss from cyber-attacks is rather small and decreasing.

The source of information on cyber-attacks in most prior studies is the firm itself. However, the firm may have strong incentives to withhold information on cyber-attacks,

especially when the occurrence of the cyber-attack and the damage it caused are uncertain. Unlike prior studies, we classify cyber-attacks into two main groups: cyber-attacks the attacked firms disclosed, and cyber-attacks that were withheld and later independently discovered by sources outside the firm. We show the market reaction to disclosed attacks is indeed small, but the market reaction to withheld attacks is negative and significant.

Using market reactions to cyber-attacks that were disclosed and cyber-attacks that were withheld and later discovered, we estimate the extent to which firms withhold information on cyber-attacks. We find managers disclose less severe attacks, and withhold information from investors on attacks that cause greater damage. The evidence is consistent with the theory that managers will not disclose negative information below a certain threshold when investors are uncertain about whether the firm possesses negative information.

The proportion of the market reaction to withheld cyber-attacks to disclosed cyber-attacks also implies managers disclose cyber-attacks only when investors already suspect that with a 46% chance an attack has occurred. When the likelihood is lower that investors will discern the existence of a cyber-attack, it is worthwhile for managers to withhold the information. Overall, our analyses suggest there is little voluntary disclosure of cyber-attacks. If regulators wish to ensure information on cyber-attacks reaches investors, they should consider imposing stricter mandatory disclosure rules regarding cyber-attacks.

References

Amir, E., and A. Ziv, 1997. Recognize, Disclose or Delay; Timing the Adoption of SFAS No. 106. *Journal of Accounting research* 35 (Spring): 61-81.

Bebchuk, L., Cohen, A., and Ferrell, A., 2009. What matters in corporate governance? *Review of Financial Studies*, 22(2), 783-827.

Campbell, K., Gordon, L., Loeb, M., and Zhou, L., 2003. The economic cost of publicly announced information security breaches: empirical evidence from the stock market. *Journal of Computer Security* 11, 431-448.

Cavusoglu, H., Mishra, B., and Raghunathan, S., 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, 69-104.

Chambers, A., Penman, S., 1984. Timeliness of reporting and the stock price reaction to earnings announcements. *Journal of Accounting Research*, 21-47.

Chernick, M. 2007. Bootstrap Methods: A Guide for Practitioners and Researchers, 2nd Edition. Wiley, New York.

Dye, R., 1985. Disclosure of nonproprietary information. *Journal of Accounting Research*, 123-145.

Ettredge, M., Richardson, V., 2003. Information transfer among Internet firms: the case of acker attacks. *Journal of Information Systems 17*, 71-82.

Fama, E., and French, K., 1996. The CAPM is wanted, dead or alive. *The Journal of Finance* 51(5), 1947-1958.

Ge, W., and McVay, S., 2005. The disclosure of material weaknesses in internal control after the Sarbanes-Oxley Act. *Accounting Horizons* 19(3), 137-158.

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

Grossman, S., 1981. The informational role of warranties and private disclosure about product quality. *Journal of Law and Economics* (December 1981), 461-83.

Grossman, S., and Hart, O., 1980. Disclosure laws and takeover bids. *Journal of Finance* (May 1980), 323-34.

Hilary, G., Segal, B., and Zhang, M., 2016. Cyber-risk disclosure: Who cares? Working paper, Georgetown University and Fordham University.

Hovav, A., D'Arcy, J., 2003. The impact of denial-of-service attack announcements on the market value of firms. *Risk Management and Insurance Review* 6. 97-121.

Jung, W., Kwon, Y., 1988. Disclosure when the market is unsure of information endowment of managers. *Journal of Accounting Research* 26 (1), 146–53.

Kannan A., Rees, J., and Shridhar, S., 2007. Market reactions to information security breach announcements: an empirical analysis, International Journal of Electronic Commerce 12, 69-91.

Kasznik, R., and Lev, B., 1995. To warn or not to warn: Management disclosures in the face of an earnings surprise. *Accounting Review*, 113-134.

Kothari, S. P., Shu, S., & Wysocki, P., 2009, Do managers withhold bad news? *Journal of Accounting Research*, 47(1), 241-276.

Kvochko, E., Pant, R., 2015. Why data breaches don't hurt stock prices. Harvard Business Review, March 31, 2015.

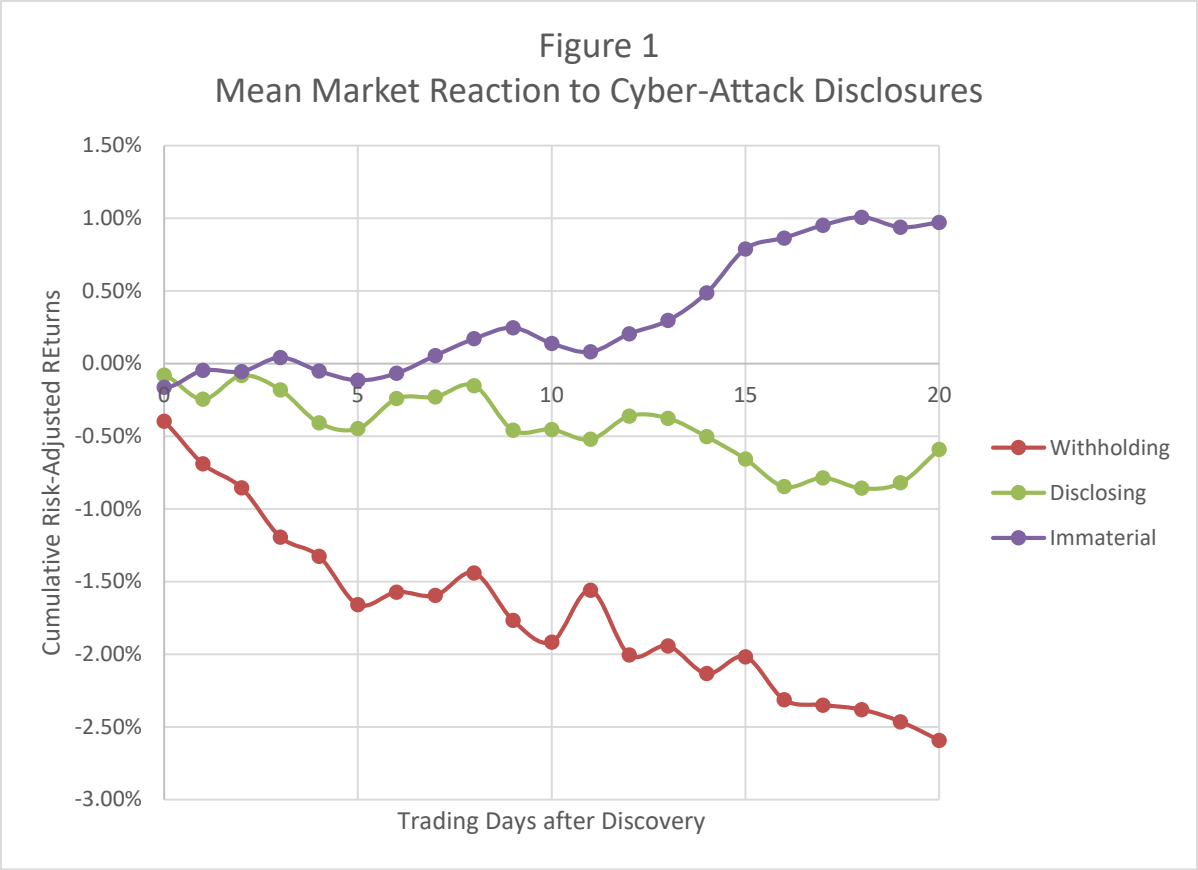
Levitt, A., 1998. The numbers game. *The CPA Journal* 68.12, 14-19.

Securities and Exchange Commission (SEC). 2011. Division of Corporation Finance, CF Disclosure Guidance, Topic No. 2 – Cybersecurity (October).

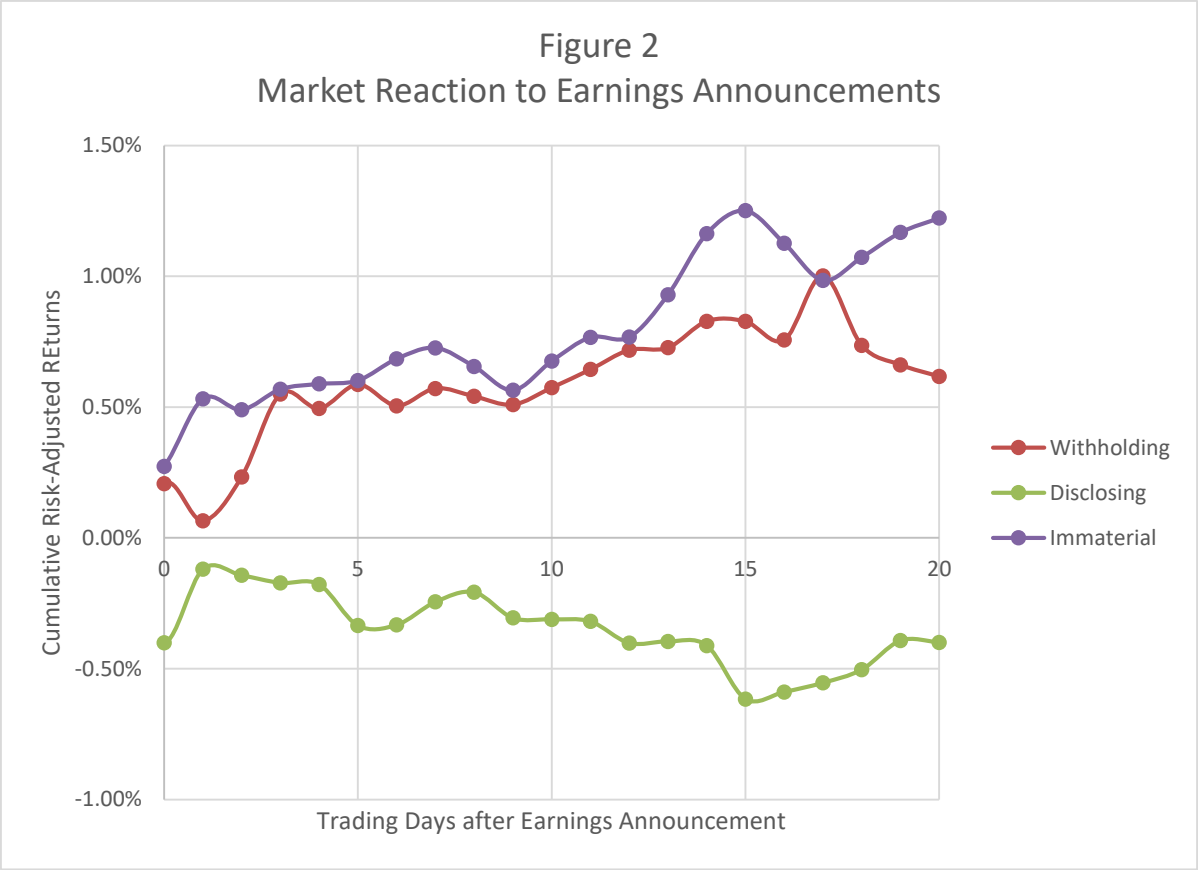
Skinner, D., 1994. Why firms voluntarily disclose bad news? *Journal of Accounting Research*, 38-60.

Skinner, D., 1997. Earnings disclosures and stockholder lawsuits. *Journal of Accounting and Economics* 23, 249-282.

Verizon Enterprise Solutions, 2015. Verizon 2015 Data Breach Investigations Report http://www.verizonenterprise.com/resources/reports/rp_data-breach-investigation-report_2015_en_xg.pdf



The figures present the stock market reaction to withholding and disclosing information on cyber-attacks. "Withholding" are cases in which the firm had not disclosed the cyber-attack for at least two days after it learned of its occurrence, and a party outside the firm consequently discovered the attack. "Disclosing" are cases in which the firm discloses the cyber-attack no later than its discovery by investors. We also present market reaction for cases without material damage ("Immaterial"); in these cases, an outsider discovered the attack, but the firm communicated that the attack caused no material damage. The cumulative risk-adjusted returns from one trading day prior to the discovery date until 20 days after the discovery date is calculated in each case, and Figure 1 presents the mean returns for stocks in each of the three portfolios. The sample includes 320 cyber-attacks between 2010 and 2015.



The graph presents a placebo test: the stock market reaction to the earnings announcements of withholding and disclosing firms on the year of the cyber-attack. For firms in the "Disclosing," "Withholding," and "Immaterial" groups, we present the cumulative risk-adjusted returns around the quarterly earnings announcements from one trading day prior to the earnings announcement date until 20 days after the announcement date. "Withholding" are cases in which the firm had not disclosed the cyber-attack for at least two days after it learned of its occurrence, and a party outside the firm consequently discovered the attack. "Disclosing" are cases in which the firm discloses the cyber-attack no later than its discovery by investors. "Immaterial" are cases in which an outsider discovered the attack, but the firm communicated that the attack caused no material damage. The sample includes 320 cyber-attacks between 2010 and 2015.

Table 1
Sample Selection

Year	Number of Cyber-Attacks	Number of Firms	Attack Type		
			Availability	Confidentiality	Integrity
2010	21	19	4	12	5
2011	35	29	6	16	12
2012	48	42	13	17	17
2013	61	48	17	32	12
2014	97	78	15	53	28
2015	58	48	13	32	13

Note: The table presents the number of cyber-attack incidents and firms in our sample by year and by type of attack. The incidents occurred between 2010 and 2015.

Table 2
Descriptive Statistics

Panel A: Disclosing vs. Withholding

Variable	(a) Disclosure			Exp. Relation	(b) Withholding			(a) – (b) *P*-values	
	#Obs.	Mean	Median		#Obs.	Mean	Median	*t*-test	*W-test*
Analysts	163	13.30	11.50	>	57	9.11	5.50	(0.01)***	(0.01)***
SOX404	152	0.07	0.00	<	55	0.56	0.00	(0.01)***	(<.01)***
Entrenchment	154	1.40	1.00	<	49	1.74	1.50	(0.04)**	(0.05)**
Hi-Tech	163	0.21	0.00	>	57	0.07	0.00	(0.01)***	(0.01)***
Damage	21	0.006	0.001	<	13	0.018	0.003	(0.10)*	(0.22)
Severity	163	4.28	3.90	<	57	4.61	5.00	(0.21)	(0.18)
ROA	163	0.05	0.04	?	57	0.05	0.04	(0.46)	(0.46)
MV	163	47,938	12,055	>	57	28,390	5,867	(0.07)*	(0.07)

Panel B: Disclosing vs. Immaterial

Variable	(a) Disclosure			Exp. Relation	(b) Immaterial			(a) – (b) *P*-values	
	#Obs.	Mean	Median		#Obs.	Mean	Median	*t*-test	*W-test*
Analysts	163	13.30	11.50	?	100	16.25	16.79	(0.05)*	(0.12)
SOX404	152	0.07	0.00	?	93	0.09	0.00	(0.39)	(0.46)
Entrenchment	154	1.40	1.00	?	94	1.27	1.00	(0.19)	(0.11)
Hi-Tech	163	0.21	0.00	?	100	0.22	0.00	(0.41)	(0.41)
Damage	21	0.006	0.001	>	6	0.000	0.000	(0.14)	(0.01)***
Severity	163	4.28	3.90	>	100	3.16	2.10	(<.01)***	(<.01)***
ROA	163	0.05	0.04	?	100	0.06	0.04	(0.05)**	(0.17)
MV	163	47,938	12,055	?	100	78,326	26,518	(<.01)***	(<.01)***

Note:

(1) The table presents characteristics of firms in three subsamples: “Disclosing,” “Withholding,” and “Immaterial.” “Disclosing” are cases in which the firm discloses the cyber-attack before an outsider discovers it. “Withholding” are cases in which the firm had not disclosed the cyber-attack for at least two days after it learned of its occurrence, and a party outside the firm consequently discovered the attack. “Immaterial” are cases in which an outsider discovered the attack, but the firm communicated that the attack caused no material damage.

(2) “Analysts” is the number of analysts covering the firm. “SOX404” is the number of material weakness in the preceding five years, where a larger value indicates weaker corporate governance quality. “Entrenchment” is an index of managers’ entrenchment, where a larger entrenchment value indicates weaker corporate governance quality. “Hi-Tech” is an indicator variable that is equal to 1 for firms in the Hi-Tech industry. “Damage” is the dollar damage disclosed by the attacked firm, divided by the market value of equity. “Severity” is an index that takes the values 0-10 based on the severity of the cyber-attack. “ROA” is net income divided by total assets in the year before the attack. “MV” is the market value of equity in millions at the beginning of the year.

(3) The sample includes 320 cyber-attacks between 2010 and 2015. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively, for the p-values reported in the parentheses.

Table 3
Firm Characteristics – Multivariate Analysis

	(1) **Disclosing vs. Withholding**		(2) **Disclosing vs. Immaterial**
Analysts	0.003		0.009
	(0.41)		(0.17)
SOX404	-0.234		-0.120
	(0.03)**		(0.23)
Entrenchment	-0.115		-0.024
	(0.10)*		(0.39)
Hi-Tech	0.684		-0.001
	(0.03)**		(0.50)
Severity	-0.061		0.141
	(0.06)*		(<.01)***
ROA	0.403		0.656
	(0.40)		(0.33)
LogMV	0.026		-0.265
	(0.36)		(<.01)***
# Observations	187		230
# Disclosure	143		143
R^2 (LRI)	9.44%		10.48%

Note:

(1) The table presents a logistic regression of the association between firm characteristics and the decisions to disclose versus withhold information on the cyber-attacks. The dependent variable equals 1 for "Disclosing" and 0 for "Withholding" and "Immaterial" cases in columns (1) and (2), respectively. "Disclosing" are cases in which the firm discloses the cyber-attack before an outsider discovers it. "Withholding" are cases in which the firm had not disclosed the cyber-attack for at least two days after it learned of its occurrence, and a party outside the firm consequently discovered the attack. "Immaterial" are cases in which an outsider discovered the attack, but the firm communicated that the attack caused no material damage.

(2) See Table 2 for variable definitions.

(3) The sample includes 320 cyber-attacks between 2010 and 2015. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. P-values are reported in the parentheses.

Table 4
Market Reaction to Cyber-Attack Disclosures

Panel A: Market reaction to withholding and disclosure

Cyber-Attack Disclosures	#Obs	Ret (-1,3)	Ret (-1,10)	Ret (-1,20)
Disclosing	163	-0.17%	-0.45%	-0.59%
		(0.28)	(0.17)	(0.19)
Withholding	57	-1.19%	-1.92%	-2.59%
		(<.01)***	(<.01)***	(<.01)***
Immaterial	100	0.04%	0.14%	0.97%
		(0.45)	(0.38)	(0.06)*

Panel B: Implied withholding probability

	Ret (-1,3)	Ret (-1,10)	Ret (-1,20)
prob(withholding)	30%	47%	46%
p-value	(<.01)***	(<.01)***	(<.01)***

Panel C: Median market reaction to withholding and disclosure

Cyber-Attack Disclosures	#Obs	Ret (-1,3)	Ret (-1,10)	Ret (-1,20)
Disclosing	163	-0.03%	-0.85%	-0.59%
		(0.21)	(0.02)**	(0.07)*
Withholding	57	-1.19%	-0.98%	-2.78%
		(<.01)***	(0.01)***	(<.01)***
Immaterial	100	-0.23%	-0.26%	0.80%
		(0.38)	(0.45)	(0.11)

Notes:

1. Panel A presents the stock market reaction to withholding and disclosing information on cyber-attacks. "Withholding" are cases in which the firm had not disclosed the cyber-attack for at least two days after it learned of its occurrence, and a party outside the firm consequently discovered the attack. "Disclosing" are cases in which the firm discloses the cyber-attack before an outsider discovers it. Finally, we present market reaction for cases without material damage ("Immaterial"); in these cases, an outsider discovered the attack, but the firm communicated that the attack caused no material damage.
2. We present cumulative risk-adjusted returns from one trading day prior to the discovery date until three, 10, and 20 days after the discovery date, labeled "Ret(-1,3)," "Ret(-1,10)," and "Ret(-1,20)", respectively. Returns are risk-adjusted using value-weighted market returns.
3. Panel B presents the implied probability of withholding, *prob(withholding)*, the cyber-attack from investors.

4. Panel C presents the median stock reaction for the Withholding, Disclosing, and Immaterial cases.

$$prob(withholding) = \frac{2 * Return\ reaction\ to\ Disclosing}{Return\ reaction\ to\ Withholding}$$

5. The sample includes 320 cyber-attacks between 2010 and 2015.
6. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively, for the p-values reported in the parentheses.

Table 5
Market Reaction after Controlling for Damage Reported by Firms

		Dependent Variable					
		Ret(-1,3)		Ret(-1,10)		Ret(-1,20)	
Independent Variables	Sign	(1)	(2)	(3)	(4)	(5)	(6)
Intercept		0.008	0.008	0.012	0.012	0.013	0.011
		(0.29)	(0.28)	(0.29)	(0.29)	(0.32)	(0.31)
Disclosing	-	-0.009	-0.007	-0.016	-0.014	-0.018	-0.015
		(0.30)	(0.34)	(0.25)	(0.28)	(0.25)	(0.28)
Withholding	-	-0.030	-0.023	-0.021	-0.014	-0.040	-0.031
		(0.05)**	(0.10)*	(0.21)	(0.30)	(0.09)*	(0.15)
Damage	-		-0.382		-0.396		-0.493
			(0.06)*		(0.13)		(0.11)
# Observations		40	40	40	40	40	40
R^2		9.76%	15.87%	1.89%	5.27%	5.53%	9.53%

Note:

1. The table presents estimation results of equation (2):
$$\mathrm{Ret}(-1,3)_i = \beta_0 + \beta_1 Disclosing_i + \beta_2 Withholding_i + \beta_3 Damage_i + \varepsilon_i .$$
2. "Ret(-1,3)" is the cumulative risk-adjusted returns from one day before to three trading days after the date the market learned of the attack. "Disclosing" is an indicator variable that equals 1 when the firm discloses the cyber-attack before an outsider discovers it. "Withholding" is an indicator variable that equals 1 when the firm had not disclosed the cyber-attack for at least two days after it learned of its occurrence, and a party outside the firm consequently discovered the attack. "Damage" is the dollar damage disclosed by the attacked firm, divided by the market value of equity.
3. The sample includes 40 firms that eventually disclosed a damage estimate.
4. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively, for the *p*-values reported in parentheses.

Table 6
Market Reaction after Controlling for Firm Characteristics

		Dependent Variable		
		Ret(-1,3)	Ret(-1,10)	Ret(-1,20)
Independent Variables	Sign			
Intercept	?	0.005	-0.067	-0.094
		(0.43)	(0.05)**	(0.05)**
Disclosing	-	-0.002	-0.003	-0.008
		(0.32)	(0.14)	(0.20)
Withholding	-	-0.013	-0.008	-0.016
		(0.01)***	(0.14)	(0.08)*
Severity	-	-0.001	-0.002	-0.003
		(0.10)*	(0.09)*	(0.10)*
Analysts	?	-0.001	-0.001	-0.001
		(0.01)***	(0.02)**	(0.12)
Hi-Tech	?	0.005	0.010	0.016
		(0.17)	(0.06)*	(0.12)
SOX404	?	-0.001	-0.006	-0.005
		(0.18)	(0.01)***	(0.06)*
Entrenchment	?	-0.003	-0.005	-0.006
		(0.13)	(0.14)	(0.23)
ROA	?	0.032	0.007	-0.019
		(0.25)	(0.46)	(0.43)
LogMV	?	0.001	0.005	0.008
		(0.36)	(0.01)***	(0.01)***
# Observations		277	277	277
R^2		5.18%	7.22%	6.28%

Note:

1. The table presents the market reaction to withholding of information on cyber-attacks, after controlling for firm characteristics:

$$Ret(-1,3) = \alpha + \beta_1 Disclosing_{it} + \beta_2 Withholding_{it} + Controls + \varepsilon_{it} \ . \qquad (3)$$

2. "Ret(-1,3)" is the cumulative risk-adjusted returns from one day before to three trading days after the date the market learned of the attack. "Disclosing" is an indicator variable

that equals 1 when the firm discloses the cyber-attack before an outsider discovers it. "Withholding" is an indicator variable that equals 1 when the firm had not disclosed the cyber-attack for at least two days after it learned of its occurrence, and a party outside the firm consequently discovered the attack.. See Table 2 for variable definitions. The regression is estimated also for the cumulative risk-adjusted returns from one trading day prior to the discovery date until 10 and 20 days after the discovery date, labeled "Ret(-1,10)" and "Ret(-1,20)," respectively.

3. The sample includes 320 observations between 2010 and 2015. Because of the lack of data on SOX404 and Entrenchment independent variables for all observations (see Table 2), the regression is estimated with 277 observations.
4. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively, for the p-values reported in parentheses, and errors are clustered by firm.

Table 7
Controlling for endogeneity

		OLS	2SLS
Independent Variables	Sign		
Intercept	?	-0.006	-0.038
		(0.42)	(0.12)
Withholding	-	-0.012	
		(0.01)***	
$\overline{\text{Withholding}}$	?		0.099
			(0.05)**
Severity	-	-0.001	-0.002
		(0.08)*	(0.03)**
Analysts	?	-0.001	-0.001
		(0.01)***	(0.01)***
Hi-Tech	?	0.004	0.008
		(0.22)	(0.08)*
SOX404	?	-0.001	-0.001
		(0.16)	(0.11)
Entrenchment	?	-0.003	-0.003
		(0.13)	(0.12)
ROA	?	0.044	0.047
		(0.19)	(0.17)
LogMV	?	0.001	0.002
		(0.22)	(0.11)
# Observations		277	277
R^2		4.98%	5.16%

Note:

1. To control for endogeneity, the fact severity of the attack can affect firms' decision to disclose or withhold, we use 2SLS estimation:

$$Withholding_{it} = \alpha + \beta HDState_{it} + \varepsilon_{it} \quad (4a)$$

$$Ret(-1,3)_{it} = \alpha + \beta_1 \overline{Withholding_{it}} + \beta_2 Severity_{it} + Controls + \varepsilon_{it} \quad . \quad (4b)$$

2. "HDState" is an indicator variable that equals 1 for high-disclosure states. "Ret(-1,3)" is the cumulative risk-adjusted returns from one day before to three trading days after the date the market learned of the attack. "Withholding" is an indicator variable that equals 1 when the firm had not disclosed the cyber-attack for at least two days after it learned

of its occurrence, and a party outside the firm consequently discovered the attack. "$\overline{\text{Withholding}}$" is the expected value from eq. (4a). See Table 2 for variable definitions.

3. The sample includes 320 observations between 2010 and 2015. Because of the lack of data on SOX404 and Entrenchment independent variables for all observations (see Table 2), the regression is estimated with 277 observations.
4. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively, for the p-values reported in parentheses, and errors are clustered by firm.

Table 8
Controlling for selection bias

		Dependent Variable	
		Withholding	Ret(-1,3)
Independent Variables	Sign		
Intercept	?	0.227 (0.41)	0.036 (0.15)
Withholding	-		-0.012 (0.01)***
Severity	-	0.102 (0.01)***	-0.008 (0.05)**
Analysts	?	0.001 (0.47)	-0.001 (0.01)***
HiTech	?	-0.702 (0.02)**	0.054 (0.07)*
SOX404	?	0.196 (0.02)**	-0.009 (0.04)**
Entrenchment	?	0.085 (0.15)	-0.011 (0.06)*
ROA	?	0,543 (0.36)	0.022 (0.32)
LogMV	?	-0.106 (0.05)**	0.008 (0.07)*
HDState		-0.103 (0.36)	
MR			-0.082 (0.08)*
# Observations		277	277
R^2		11.21%	5.57%

Note:

1. To control for selection bias, the fact discovered attacks may be a biased sample of the withheld attacks, we use Heckman's correction:

$$Withholding_i = \delta_0 + \delta_1 HDState_i + \delta_2 Severity_i + \delta_3 Analysts_i + \delta_4 HiTech_i + \delta_6 SOX\,404_i + \delta_7 Entrenchment_i + \delta_8 ROA_i + \delta_9 LogMV_i + \varepsilon_i \quad (5a)$$

$$Ret(-1,3)_{it} = \alpha + \beta_1 Withholding_{it} + \beta_2 Severity_{it} + Controls + MR + \theta_{it} \quad (5b)$$

1. "HDState" is an indicator variable that equals 1 for high-disclosure states. "Ret(-1,3)" is the cumulative risk-adjusted returns from one day before to three trading days after the date the market learned of the attack. "Withholding" is an indicator variable that equals 1 when the firm had not disclosed the cyber-attack for at least two days after it learned of its occurrence, and a party outside the firm consequently discovered the attack. "MR" is the inverse Mills ratio based on eq. (5a). See Table 2 for variable definitions.
2. The sample includes 320 observations between 2010 and 2015. Because of the lack of data on SOX404 and Entrenchment independent variables for all observations (see Table 2), the regression is estimated with 277 observations.
3. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively, for the p-values reported in parentheses, and errors are clustered by firm.

Table 9
Market Reaction with Alternative Risk Adjustments

Panel A: Size-adjusted returns

Cyber-Attack Disclosures	#Obs	Ret (-1,3)	Ret (-1,10)	Ret (-1,20)
Disclosing	136	-0.35% (0.14)	-0.68% (0.11)	-0.72% (0.16)
Withholding	50	-1.18% (<.01)***	-2.03% (<.01)***	-2.63% (<.01)***
Immaterial	91	0.01% (0.44)	0.14% (0.39)	1.21% (0.03)**

Panel B: Returns adjusted for size, book-to-market, and momentum

Cyber-Attack Disclosures	#Obs	Ret (-1,3)	Ret (-1,10)	Ret (-1,20)
Disclosing	117	-0.36% (0.14)	-0.52% (0.16)	-0.37% (0.30)
Withholding	42	-1.19% (<.01)***	-2.14% (<.01)***	-2.55% (<.01)***
Immaterial	81	0.18% (0.29)	0.16% (0.36)	1.17% (0.04)**

Notes:

1. Panel A presents the stock market reaction with size-adjusted returns, and Panel B with stock returns adjusted for risk using size, book-to-market, and momentum portfolios. We present cumulative risk-adjusted returns from one trading day prior to the discovery date until three, 10, and 20 days after the discovery date, labeled "Ret(-1,3)," "Ret(-1,10)," and "Ret(-1,20)," respectively. "Withholding" are cases in which the firm had not disclosed the cyber-attack for at least two days after it learned of its occurrence, and a party outside the firm consequently discovered the attack. "Disclosing" are cases in which the firm discloses the cyber-attack before an outsider discovers it. "Immaterial" are cases in which an outsider discovered the attack, but the firm communicated that the attack caused no material damage.
2. The sample includes 320 cyber-attacks between 2010 and 2015.
3. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively, for the p-values reported in the parentheses.

Table 10
Market Reaction on Earnings Announcements

Independent Variables	Sign	Ret(-1,3)
Intercept	?	-0.079 (0.16)
Disclosing	?	-0.035 (0.30)
Withholding	?	0.003 (0.38)
Severity	?	-0.0004 (0.34)
Analysts	?	-0.0003 (0.16)
Hi-Tech	?	0.010 (0.20)
SOX404	?	-0.002 (0.38)
Entrenchment	?	0.005 (0.09)*
ROA	?	0.074 (0.16)
LogMV	?	0.004 (0.15)
Observations		1,043
R^2		1.98%

Note:

1. The table presents the market reaction to the earnings announcements on the year of the cyber-attacks:

$$Ret(-1,3) = \alpha + \beta_1 Disclosing_{it} + \beta_2 Withholding_{it} + Controls + \varepsilon_{it} \ . \qquad (3)$$

2. "Ret(-1,3)" is the cumulative risk-adjusted returns from one day before to three trading days after the date the market learned of the attack. "Disclosing" is an indicator variable that equals 1 when firms disclose the cyber-attack before an outsider discovers it. "Withholding" is an indicator variable that equals 1 when the firm had not disclosed the cyber-attack for at least two days after it learned of its occurrence, and a party outside the firm consequently discovered the attack. See Table 2 for variable definitions. The regression is estimated also for the cumulative risk-adjusted returns from one trading day prior to the discovery date until 10 and 20 days after the discovery date, labeled "Ret(-1,10)" and "Ret(-1,20)," respectively.
3. The sample includes 1,043 quarterly earnings announcements between 2010 and 2015.
4. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively, for the p-values reported in parentheses, and errors are clustered by firm.